Mª TERESA PADILLA

TÉCNICAS E INSTRUMENTOS PARA EL DIAGNÓSTICO Y LA EVALUACIÓN EDUCATIVA

EDITORIAL CCS

Mª Teresa PADILLA CARMONA

TÉCNICAS E INSTRUMENTOS PARA EL DIAGNÓSTICO Y LA EVALUACIÓN EDUCATIVA

Editorial CCS

Mª Teresa Padilla Carmona es Doctora en Ciencias de la Educación y profesora de la Universidad de Sevilla, en la que imparte, entre otras, la materia de *Técnicas e Instrumentos de Diagnóstico* de la *Licenciatura en Pedagogía*. Es también miembro del Laboratorio de Diagnóstico en Educación de la Facultad de Ciencias de la Educación de esa Universidad.

Su actividad investigadora se centra en la construcción y validación de estrategias diagnósticas, especialmente dentro del campo de la orientación profesional. Ha publicado varios libros entre los que cabe destacar el denominado *Estrategias para el diagnóstico y la orientación profesional de personas adultas* (2001). Asimismo, es autora y coautora de numerosos artículos publicados en revistas nacionales e internacionales.

Segunda edición: noviembre 2011

Página web de EDITORIAL CCS: www.editorialccs.com

© Mª Teresa PADILLA
© 2002. EDITORIAL CCS, Alcalá, 166 / 28028 MADRID

Cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra solo puede ser realizada con la autorización de sus titulares, salvo excepción prevista por la ley. Diríjase a CEDRO (Centro Español de Derechos Reprográficos, www.cedro.org) si necesita fotocopiar o escanear algún fragmento de esta obra.

Diseño de portada: Olga R. Gambarte
Composición digital: Safekat
ISBN (epub): 978-84-9023-659-8

Índice

Portada

Créditos

Prólogo

Introducción

CAPÍTULO 1. INTRODUCCIÓN AL DIAGNÓSTICO Y LA EVALUACIÓN EN EDUCACIÓN
1. INTRODUCCIÓN AL CONCEPTO DE DIAGNÓSTICO Y EVALUACIÓN
2. OBJETIVOS, FUNCIONES Y ÁMBITOS DEL DIAGNÓSTICO EN EDUCACIÓN
3. ENFOQUES DE DIAGNÓSTICO
    3.1. Modelos tradicionales
    3.2. Modelos conductuales
    3.3. Modelos derivados de la psicología cognitiva
4. APORTACIONES DESDE LA INTERVENCIÓN ORIENTADORA: PRINCIPIOS Y MODELOS PARA EL DIAGNÓSTICO EDUCATIVO
5. FASES EN EL PROCESO DE DIAGNÓSTICO
6. EL INFORME DE DIAGNÓSTICO
    6.1. Cuestiones previas: formación, habilidad y ética en la elaboración del informe diagnóstico
    6.2. Propuesta de estructura del informe
    6.3. Recomendaciones adicionales en la elaboración de informes
    6.4. Posibles criterios para valorar informes diagnósticos
7. GARANTÍAS ÉTICAS Y DEONTOLOGÍA PROFESIONAL EN EL DIAGNÓSTICO Y LA EVALUACIÓN
    7.1. Aspectos de la profesión pedagógica a incluir en un código deontológico
8. CONSIDERACIONES SOBRE LA SELECCIÓN DE INSTRUMENTOS PARA EL DIAGNÓSTICO Y LA EVALUACIÓN
Referencias bibliográficas

CAPÍTULO 2. INTRODUCCIÓN AL DISEÑO Y APLICACIÓN DE PROCEDIMIENTOS OBSERVACIONALES
1. LA OBSERVACIÓN COMO ESTRATEGIA DIAGNÓSTICA

1.1. Concepto y características
1.2. Modalidades de observación
1.2.1. Observación sistematizada/no sistematizada
1.2.2. Observación participante/no participante
1.2.3. Auto-observación/hetero-observación
2. PREPARACIÓN DE LA OBSERVACIÓN: LA SELECCIÓN DE MUESTRAS
3. DISEÑO Y APLICACIÓN DE PROCEDIMIENTOS OBSERVACIONALES
3.1. Observación sistematizada
3.1.1. Aspectos a tener en cuenta en la elaboración de sistemas categoriales
3.1.2. Proceso de elaboración de sistemas categoriales
3.1.3. Tipos de sistemas categoriales
3.2. Observación no sistematizada
3.2.1. Tipos de registros de observación no sistematizada
3.2.2. Registros *in situ*
3.2.3. Registros *a posteriori*: el diario
3.3. Los sistemas tecnológicos en la observación sistematizada y no sistematizada
4. GARANTÍAS CIENTÍFICAS DE LAS TÉCNICAS Y REGISTROS OBSERVACIONALES
4.1. Sesgos que afectan a las técnicas y registros observacionales
4.2. Fiabilidad y validez de los registros observacionales
4.3. Criterios de calidad alternativos para los registros de observación no sistematizados
Referencias bibliográficas

CAPÍTULO 3. INTRODUCCIÓN AL DISEÑO Y APLICACIÓN DE LAS TÉCNICAS E INSTRUMENTOS DE ENCUESTA
1. LA ENCUESTA COMO ESTRATEGIA DIAGNÓSTICA
2. LA ENTREVISTA EN EL DIAGNÓSTICO Y LA EVALUACIÓN
2.1. Entrevistas abiertas o no estructuradas
2.1.1. Preparación de la entrevista no estructurada
2.1.2. Desarrollo y aplicación de la entrevista no estructurada
2.1.3. El registro de la información en la entrevista abierta
2.2. Entrevistas estructuradas
2.2.1. Diferencias entre la entrevista estructurada y el cuestionario
2.2.2. Preparación de la cédula
2.2.3. Aplicación de la entrevista estructurada y el registro de la información
3. EL CUESTIONARIO EN LOS PROCESOS DIAGNÓSTICOS
3.1. Concepto y características
3.2. Diseño de cuestionarios y entrevistas estructuradas
3.2.1. Proceso de diseño y preparación de cuestionarios y entrevistas estructuradas
3.2.2. Las preguntas en cuestionarios y entrevistas estructuradas: tipos, selección y formulación
3.2.3. Tendencias de respuesta que inducen al sesgo
3.2.4. Preparación del formato del cuestionario
4. GARANTÍAS CIENTÍFICAS DE LAS TÉCNICAS E INSTRUMENTOS BASADOS EN LA ENCUESTA

4.1. Cuestionarios
4.2. Entrevistas estructuradas
4.3. Entrevistas abiertas
Referencias bibliográficas

CAPÍTULO 4. INTRODUCCIÓN AL DISEÑO Y APLICACIÓN DE TESTS Y PRUEBAS OBJETIVAS
1. LA MEDIDA EN EL DIAGNÓSTICO Y LA EVALUACIÓN
1.1. La medición, la evaluación y el diagnóstico
2. LOS TESTS: CONCEPTO Y CARACTERÍSTICAS
3. TIPOS DE TESTS
3.1. Algunas clasificaciones de tests aplicables al campo educativo
3.2. Clasificación temática de algunos tests y pruebas para la evaluación del desarrollo personal y profesional
3.2.1. Pruebas estandarizadas para la evaluación de los procesos de estudio y aprendizaje
3.2.2. Pruebas para la evaluación de los intereses profesionales
3.2.3. Instrumentos para la evaluación del desarrollo vocacional y profesional y la toma de decisiones
3.2.4. Instrumentos para la evaluación de los valores personales y laborales
3.2.5. Instrumentos para la evaluación de ambientes, situaciones y problemas personales y profesionales
3.2.6. Instrumentos para la evaluación del autoconcepto y la autoestima
3.2.7. Instrumentos para la evaluación de las habilidades sociales y el autocontrol
3.2.8. Instrumentos complementarios en el proceso de orientación personal y profesional
4. CONSTRUCCIÓN DE PRUEBAS OBJETIVAS
4.1. Construcción de los planos de la prueba: objetivos, contenidos y tabla de especificaciones
4.2. Formulación y selección de elementos
4.3. Ordenación de los elementos y preparación del formato de la prueba
4.4. Valoración y puntuación
5. APLICACIÓN DE TESTS: PRECAUCIONES Y COMPETENCIAS
6. INTERPRETACIÓN DE LAS PUNTUACIONES DE LOS TESTS
6.1. Puntuaciones percentiles
6.2. Puntuaciones típicas o *z* y puntuaciones típicas derivadas
7. GARANTÍAS TÉCNICAS Y CIENTÍFICAS EN LOS INSTRUMENTOS DE MEDIDA
7.1. Tendencias de respuesta que afectan a la medida
7.2. Análisis de ítems
7.3. Características técnicas de los instrumentos de medida
7.3.1. Fiabilidad
7.3.2. Validez
Referencias bibliográficas

CAPÍTULO 5. OTRAS TÉCNICAS Y ESTRATEGIAS PARA EL DIAGNÓSTICO EN EDUCACIÓN
1. INTRODUCCIÓN

2. EL PORTAFOLIOS
- 2.1. Portafolios de aprendizaje
- 2.2. Propósitos de los portafolios de aprendizaje
- 2.3. Las metas de aprendizaje
- 2.4. Contenidos del portafolios
- 2.5. Las tareas de aprendizaje
- 2.6. Modelos de portafolios de aprendizaje
- 2.7. Materiales de los portafolios de aprendizaje
- 2.8. Criterios de evaluación
- 2.9. Validez de los portafolios

3. LA HISTORIA DE VIDA
- 3.1. Emergencia histórica de la historia de vida en sus modalidades biográfica, autobiográfica y de relatos de vida
- 3.2. La historia de vida como instrumento y método de diagnóstico e intervención educativa. Presupuestos fundamentales
- 3.3. Operativización de la historia de vida. Análisis de un ejemplo: *Trozos de mi vida*

4. EL GRUPO DE DISCUSIÓN
- 4.1. El proceso de preparación y realización del grupo de discusión
- 4.2. El rol del moderador del grupo de discusión
- 4.3. Consideraciones finales

5. LAS REJILLAS DE CONSTRUCTOS PERSONALES
- 5.1. Preparación, aplicación e interpretación de la técnica de rejilla de constructos personales
- 5.2. Un ejemplo de rejilla de constructos personales aplicada a la decisión vocacional

Referencias bibliográficas

ANEXOS
Anexo 1. CÓDIGO DEONTOLÓGICO PROFESIONAL
Anexo 2. PROPUESTA DE NORMAS ÉTICASPARA LA ELABORACIÓN DE UN CÓDIGO DEONTOLÓGICO DEL PROFESIONAL DE LA PEDAGOGÍA (Versión 2001)
Anexo 3. EJEMPLOS DE SISTEMAS CATEGORIALES DE OBSERVACIÓN
Anexo 4. EJEMPLOS DE APLICACIONESDE LAS TÉCNICAS DE ENCUESTA

Notas

Prólogo

Nos encontramos ante un libro de enorme interés tanto por sus fundamentos teóricos como por su alcance práctico. Y con una doble dimensión: es útil al alumnado y también al profesorado de pedagogía, psicopedagogía y psicología. Muestra de una manera sencilla e inteligible los conceptos básicos de diagnóstico y de evaluación aplicados a la educación y, a partir de la fundamentación conceptual, se dispone a explicar a los profesionales todo un abanico de instrumentos necesarios en la práctica cotidiana.

Hablar de las técnicas de observación, de los procedimientos más idóneos de la encuesta, de los programas de tests y de pruebas objetivas y de técnicas no tan estandarizadas pero enormemente útiles como son la rejilla, el portafolios, etc., es algo que se esperaba desde hace tiempo.

María Teresa Padilla y un conjunto de colegas suyos de la Universidad de Sevilla —Isabel López Górriz, Emilia Moreno Sánchez y Eduardo García Jiménez—, preocupados por hacer extensivos los últimos hallazgos y tendencias del diagnóstico en educación y en orientación escolar y profesional (no sólo para personas escolarizadas sino también para personas adultas) se han dedicado en estos últimos años a profundizar en temáticas tan necesarias y tan complejas como son la evaluación y el diagnóstico. Y es de agradecer.

Algo muy importante en este libro es la difusión del código deontológico profesional y la propuesta de normas éticas a la hora de elaborarlo. En nuestro país faltaba de alguna manera una ejemplificación de este aspecto tan relevante de la práctica profesional intra y extraescuela.

Desde estas páginas invitamos a los nuevos profesionales de la orientación, del consejo y de la formación continua a estudiar a fondo los temas y a poner en práctica (siempre con espíritu crítico y constructivo) los instrumentos que nos señalan... Y a indagar sobre la extensa, rica y actualizada bibliografía que proponen.

El puente entre la pedagogía y la psicología hace tiempo que está tendido. Ha llegado la hora de poner manos a la obra.

María Luisa Rodríguez Moreno
CATEDRÁTICA DE ORIENTACIÓN VOCACIONAL Y FORMACIÓN PROFESIONAL

Introducción

Técnicas e instrumentos para el diagnóstico y la evaluación educativa es un libro que nace con dos importantes ambiciones. Una de ellas, la principal, la que ha estado presente cada vez que nos sentábamos ante el ordenador y seleccionábamos, una a una, las palabras con las que íbamos construyendo esta obra, es la de servir de material didáctico para el trabajo en la asignatura que impartimos actualmente, denominada *Técnicas e instrumentos de diagnóstico* y ubicada dentro de la titulación de Pedagogía de la Universidad de Sevilla. La otra finalidad es que pueda ser de utilidad para todos aquellos profesionales que se dedican al diagnóstico en educación y en disciplinas afines.

Sin embargo, ambas finalidades, en ocasiones, entran en conflicto. Así, el hecho de que éste sea un texto dirigido a los estudiantes, y diseñado para el contexto concreto de una asignatura, ha impuesto una serie de condicionantes en lo que respecta a su estructura y contenido. De esta forma, un profesional del diagnóstico educativo que lea el libro puede encontrar carencias significativas para su práctica cotidiana, lagunas que probablemente derivan de las opciones concretas que hemos adoptado.

A la hora de diseñar el programa de nuestra asignatura, hemos de tener en cuenta por un lado el perfil de la titulación en la que ésta se encuentra y, por otro, otras materias afines tanto de la misma titulación, como de otras titulaciones de las Ciencias de la Educación. Es evidente que debemos intentar no repetir los mismos contenidos, sino que cada materia debería profundizar en los aspectos concretos del saber que resultan más relevantes para su campo de estudio. De esta forma, desde hace ya varios años, otras materias, obligatorias y optativas, se ocupan de profundizar en campos concretos del diagnóstico educativo como, por ejemplo, las dificultades de aprendizaje, las necesidades educativas especiales, el diagnóstico del desarrollo cognitivo, la evaluación de programas, la validación de instrumentos, etc. Bajo estos acuerdos docentes compartidos por el profesorado, la materia de *Técnicas e instrumentos de diagnóstico* aspira, esencialmente, a promover que el estudiante sepa diseñar y aplicar técnicas e instrumentos de recogida de información, así como adoptar una actitud crítica y reflexiva ante los procedimientos que usa. Asimismo, se procura también aplicar estos conocimientos a la evaluación para el desarrollo personal y profesional, abarcando algunos aspectos como la evaluación de las actitudes, los valores, los intereses, etc. Es decir, el contexto de aplicación de las técnicas e instrumentos es la dimensión afectiva del ser humano, entendida en un sentido amplio.

La razón para citar aquí las finalidades que perseguimos con la materia y los contenidos que derivan de estas finalidades, no es otra que la de poder justificar las importantes limitaciones que este libro puede tener. Lo que, dicho de otra forma, supone

que hemos confeccionado el libro «a la medida» de nuestro trabajo docente. En todo momento, hemos querido disponer de un material que se aproxime lo mejor posible a la forma en que impartimos la materia y con la que trabajamos en el aula. Así, puesto que siempre intentamos que el alumno haga prácticas relacionadas con las técnicas de recogida de información, en la obra se ha hecho mucho hincapié en los aspectos procedimentales. De esta forma, no se trata sólo de que los estudiantes sepan en qué consiste la observación, sino también que sean capaces de diseñar y aplicar instrumentos de observación, así como de reflexionar sobre la estrategia diagnóstica que ponen en práctica. Y lo mismo para el resto de técnicas de diagnóstico.

Además de las finalidades y contenidos de la materia que impartimos, otro importante factor justifica algunas de las decisiones tomadas en cuanto a los contenidos del libro. Ese factor es el tiempo, eterno enemigo de los programas universitarios. Así, es evidente que existen muchos otros procedimientos de diagnóstico que aquí ni se han mencionado, pero también es evidente que resulta imposible abordar sistemáticamente todos ellos, siendo necesario elegir un conjunto más limitado y trabajar con más profundidad en el mismo.

Teniendo todo esto presente, hemos confeccionado un texto que se divide esencialmente en cinco grandes bloques o capítulos. El primero de ellos pretende constituirse en una introducción general al concepto y proceso de diagnóstico. Antes de adentrarnos en las distintas estrategias diagnósticas, hemos considerado necesario explorar los rudimentos de esta materia y las cuestiones básicas que permiten encuadrar las diversas técnicas en el proceso completo y complejo del diagnóstico educativo.

El segundo capítulo se dedica a la observación. En él se intenta conceptualizar en qué consiste esta técnica, presentar las distintas modalidades de observación y el proceso de planificación y preparación de la misma. También se analizan diferentes tipos de registros observacionales, aportando información sobre su diseño y aplicación, así como algunas nociones básicas sobre las garantías científicas que toda observación debe cumplir.

Las técnicas de encuesta, como el cuestionario y la entrevista (tanto abierta como estructurada), se estudian en el tercer capítulo. Nuevamente, se intenta conceptualizar cada técnica, así como introducir al lector en el diseño, aplicación y validación de la misma.

El capítulo cuarto se centra en los tests e instrumentos de medida. Para ello se intenta clarificar qué entendemos por medición, cuáles son los instrumentos que permiten tal medida, cómo se construyen algunos de estos instrumentos, qué aspectos hay que tener en cuenta para su selección y aplicación, y cuáles son los requisitos técnicos y científicos que este grupo de instrumentos debe reunir.

El capítulo quinto, con el que se cierra el libro, se ha dedicado al análisis de otros procedimientos diagnósticos difícilmente clasificables en las categorías anteriores (observación, encuesta, medida). En este capítulo han participado, junto con la autora,

Eduardo García Jiménez, Isabel López Górriz y Emilia Moreno Sánchez, los cuales han escrito, respectivamente, sobre el portafolios, la historia de vida y los grupos de discusión, estando elaborado el apartado sobre rejillas de constructos personales por la autora del libro. Todas estas estrategias, potencialmente muy atractivas para el diagnóstico educativo, sólo son un reducido conjunto de las que se podrían haber incluido en el capítulo. Sin embargo, tomando de nuevo la referencia de nuestro trabajo en la asignatura, hemos seleccionado aquéllas que, dentro del marco temporal con el que trabajamos, solemos impartir.

Por último, queremos aprovechar esta introducción para agradecer las múltiples colaboraciones, directas e indirectas, de las que nos hemos beneficiado. De forma indirecta, agradecemos a muchas y muchos estudiantes de pedagogía las preguntas y comentarios que nos hacen en clase y en los pasillos. No saben ellos hasta qué punto estas intervenciones nos obligan a pensar, a matizar nuestras afirmaciones, a mejorar y ampliar nuestros conocimientos. De forma indirecta queremos agradecer también a los autores de tantos libros de cuya sabiduría y experiencia nos hemos aprovechado para escribir éste; y no sólo nos hemos alimentado de autores de libros, sino también de compañeros de trabajo, como Isabel López, que, además de su apoyo, nos han transmitido parte de sus conocimientos y su saber.

De forma directa, tenemos también algunas deudas importantes. A Eduardo García le debemos los apuntes con los que empezamos a impartir la asignatura. Desde entonces la materia ha evolucionado, pero no cabe ninguna duda de que aquellos apuntes nos han "estructurado" el pensamiento y, por tanto, de alguna forma, fueron la *proto-materia* de este libro. A Mª Carmen Pastor, Magdalena Suárez y Guillermo Castrejón les tenemos que agradecer sus correcciones y sugerencias, lo cual ha aliviado y aligerado nuestro propio trabajo que hoy, por fin, podemos dar por concluido... hasta que llegue el momento de más correcciones, claro.

CAPÍTULO

1

INTRODUCCIÓN AL DIAGNÓSTICO Y LA EVALUACIÓN EN EDUCACIÓN

1. INTRODUCCIÓN AL CONCEPTO DE DIAGNÓSTICO Y EVALUACIÓN

El término *diagnóstico* proviene de las palabras griegas *dia* (a través) y *gignosco* (conocer). Por esta razón, su significado viene a expresar una forma de conocimiento obtenida a través de indicios, como la que realizan los profesionales de la medicina, los cuales observan una serie de síntomas, y a partir de ellos deducen alguna anomalía o enfermedad. De hecho, el diccionario recoge la palabra *diagnóstico* como propia del campo profesional de la medicina y le otorga el significado de «conjunto de signos que sirven para fijar el carácter de una enfermedad».

Sin embargo, el diagnóstico también ha tenido una fuerte vinculación con el campo de la psicología aplicada. En 1921 Rorschach introduce el término *psicodiagnóstico* para referirse al campo específico de las enfermedades mentales. Se trata, nuevamente, de determinar enfermedades, sin embargo, estas anomalías no son ahora necesariamente físicas o relacionadas con la salud corporal, sino de carácter psicológico.

Esta vinculación inicial del diagnóstico en educación con la medicina y la psicología, configura una determinada concepción de esta actividad. Así, en el campo educativo, el diagnóstico se enfoca a detectar «anomalías» en el proceso escolar. Ello tiene algunas implicaciones entre las que podemos destacar las siguientes:

— El diagnóstico es una actividad independiente y desligada de la intervención educativa.

— Se aplica sólo a determinados alumnos y alumnas: aquellos que detentan una «enfermedad» (dificultad o problema de aprendizaje).

— La intervención diagnóstica, al igual que la médica, está descontextualizada. No se vincula al contexto de aprendizaje en el que se producen las dificultades.

Puede, por tanto, decirse que la introducción del término *diagnóstico* al campo educativo se hizo desde esta acepción médica, que implica abordar a los sujetos con dificultades desde modelos clínicos. Esta concepción del diagnóstico que, por el momento, podemos denominar de forma genérica como enfoque tradicional, está hasta cierto punto desfasada y superada por otros enfoques y creencias que, cada vez más, vienen a poner el acento en la profunda relación que debe existir entre el diagnóstico y la intervención psicopedagógica (BUISÁN y MARÍN, 1987; PÉREZ JUSTE, 1990; SANZ ORO, 1990; SOBRADO y OCAMPO, 1997; etc.), llegando algunos de estos autores (GARCÍA NIETO, 1990) a identificar ambos conceptos.

Así, el diagnóstico aplicado a la educación ha ido ampliando su campo. Se han dado nuevas definiciones de esta actividad y se han desarrollado nuevas aplicaciones de la misma. Ya no se hace referencia en exclusividad a las características de los educandos,

sino también a los *agentes y contextos educativos*. Asimismo, la tradicional función correctiva del diagnóstico se ha ampliado con dimensiones más *preventivas* y de *desarrollo*. Algunas notas características de estos enfoques más actuales son:

— El diagnóstico es una actividad que se aplica a todo el alumnado. En principio, todos los alumnos y las alumnas son especiales y, por lo tanto, la intervención educativa debe adaptarse a todos ellos.

— La función del diagnóstico no es la clasificación de los sujetos, sino que su carácter es esencialmente formativo, de evaluación continua, para favorecer la toma de decisiones y la mejora de los procesos y productos educativos.

— En lugar de primar el carácter terapéutico y correctivo (detectar y solucionar problemas), el diagnóstico ha de ser fundamentalmente preventivo y de desarrollo de las potencialidades.

— Se ha de realizar dentro de los contextos en los que interactúa el alumnado. De hecho, estos contextos serán también objeto de evaluación.

Aunque posteriormente profundizaremos en los enfoques del diagnóstico, podemos apreciar cómo se ha producido un cambio en el sujeto del mismo. Así, desde una concepción tradicional, sólo es sujeto del diagnóstico el alumno con problemas, mientras que desde las perspectivas más actuales, el diagnóstico no sólo se aplica a todos los alumnos y alumnas, sino que también afecta a los centros, los contextos, los agentes y los procesos educativos.

Este cambio en el enfoque implica también un cambio terminológico. Durante mucho tiempo ha existido cierta polémica en torno a los términos *diagnóstico* y *evaluación*, encontrándose autores que defienden el uso diferenciado de ambos y otros que se muestran más partidarios de usar indistintamente los dos términos.

Los especialistas que defienden la diferenciación de ambos conceptos puntualizan que *diagnóstico* es el término en castellano correspondiente a *assessment* (valoración), que se utiliza para indicar su carácter descriptivo y su aplicación a sujetos. *Evaluación* sería equivalente al inglés *evaluation*, que se aplica más específicamente a entidades abstractas como programas, currículos, centros e instituciones. A diferencia del diagnóstico, la evaluación está enfocada a la toma de decisiones y la formulación de juicios de valor.

En la práctica del diagnóstico en el campo escolar también ha existido tradicionalmente una diferenciación entre *diagnosticar* y *evaluar:* los sujetos a quienes se aplican, las variables en las que se centran, los agentes que los realizan, etc. Así, el diagnóstico está vinculado a *variables* psicobiológicas, tales como las aptitudes, los intereses o la personalidad (obsérvese la influencia de la corriente médico-psicológica), mientras que la evaluación se liga fundamentalmente a los aprendizajes, con especial incidencia en ámbitos académicos y preferentemente en el dominio cognoscitivo (conocimientos).

Por otra parte, también los *sujetos* a quienes se aplican suelen ser distintos (PÉREZ JUSTE, 1994). En el diagnóstico se piensa inicialmente en aquellos que tienen problemas o dificultades (nuevamente nos aparece la influencia de la corriente médico-psicológica: «descripción de síntomas para fijar una enfermedad»), mientras que en la evaluación se trabaja con alumnado «normal». Así, «el diagnóstico se concentra en "descubrir e interpretar signos", la evaluación lo hace en "valorar, justipreciar, tasar" algo que, por lo general, se traduce en marcos académicos en términos de calificación» (PÉREZ JUSTE, 1994:589).

Pueden encontrarse también diferencias en los *momentos* en que se aplican (PÉREZ JUSTE, 1994): el diagnóstico suele preceder a la actuación de los profesionales de la educación, ya que se concibe como orientación y guía del tratamiento; la evaluación, por su parte, se concentra en la etapa posterior al tratamiento, para comprobar el progreso y resultados del mismo.

Por último, son igualmente diferentes los *agentes* en ambos casos: el diagnóstico es competencia de profesionales especializados que se hallan, en muchos casos, fuera del sistema educativo o del centro escolar; por contra, la evaluación es una de las funciones habituales del profesorado.

A pesar de esta aparente claridad en la diferenciación de ambos conceptos, se están dando un conjunto de circunstancias que llevan a una cada vez mayor similitud entre estos dos campos, resultando difícil diferenciar hoy día sus respectivas competencias (GARCÍA JIMÉNEZ, 1994; PÉREZ JUSTE, 1994):

— La intervención de los especialistas no está hoy día dirigida en exclusividad al alumnado con problemas, sino a todos los educandos. Todo el alumnado es especial y, por tanto, la educación debe adecuarse a todos ellos (educación individualizada). Esto imprime al diagnóstico un carácter preventivo y de desarrollo de las potencialidades, que excede la mera función terapéutica y correctiva que esta actividad ha tenido en sus orígenes.

— La evaluación pierde, cada vez más, su función acreditadora/sancionadora, para hacerse formativa y continua. Esto sobreviene como consecuencia de desligarnos de los planteamientos médico-psicológicos y comenzar a pensar en el diagnóstico/evaluación desde el plano educativo: no nos interesan aquí sólo los resultados, sino también los procesos que necesitan ser mejorados. Ya no se evalúa únicamente al alumnado, sino también a los programas, metodologías, profesorado, centros e instituciones.

— Tanto el diagnóstico como la evaluación no tienen sentido por sí mismos, sino dentro de una propuesta de intervención educativa. Cuando diagnosticamos o evaluamos perseguimos la mejora de una situación, por lo que toda intervención educativa que no tenga en su origen un análisis de necesidades (por informal que éste sea) puede convertirse en la acción por la acción. Al mismo tiempo, el diagnóstico/evaluación debe estar enfocado a la intervención y la mejora.

— Por otra parte, el diagnóstico se ha visto obligado a modificar su centro de atención desde el «sujeto» al «sujeto y su entorno». El diagnóstico ha de referir sus explicaciones y pronósticos a las exigencias sociales y escolares que se explicitan en un contexto educativo concreto. Por esta razón, cada vez es más necesaria la participación en el diagnóstico/evaluación de los diversos agentes educativos (profesionales de la orientación, padres y madres, profesorado, alumnado...).

A la luz de estos cambios en la forma de entender la actividad educativa, el diagnóstico y la evaluación en educación, parece justificada, por tanto, la cada vez mayor similitud existente entre ambos términos y el hecho de que hoy se les considere como actividades con contenidos similares (GARCÍA JIMÉNEZ, 1994).

No obstante, existen diferentes posicionamientos a la hora de responder a la pregunta *«¿es lo mismo diagnosticar y evaluar?»*. Así, encontramos autores que postulan una diferencia etimológica y conceptual entre ambos. Éste es el caso de CASTILLO (1994), para quien el *diagnóstico* tiene lugar en el momento inicial de todo proceso didáctico y está enfocado a describir y explorar, mientras que la *evaluación* supone el seguimiento crítico del proceso y está encaminada a establecer juicios de valor con intencionalidad perfectiva y sancionadora. A pesar de ello, admite la cooperación y complementariedad entre ambos procesos e incluso la existencia de elementos comunes a los mismos (por ejemplo, los resultados de una evaluación contribuyen al diagnóstico, mientras que los datos aportados por un diagnóstico facilitan la evaluación). Sin embargo, en opinión de este autor, sus fines difieren, ya que la *evaluación* sirve para determinar en qué grado se lograron los objetivos propuestos, mientras que el *diagnóstico* permite descubrir y explorar dificultades que pueden interferir en el proceso de enseñanza-aprendizaje.

Otros autores, más partidarios de la similitud que de la diferencia, centran la diferencia entre ambos conceptos en la cuestión del alcance del diagnóstico respecto a la evaluación. En nuestro contexto, cabe mencionar dos autores que responden de forma diferente a esta cuestión. Por un lado, LÁZARO (1986, 1990) defiende que el proceso de diagnóstico no concluye con la valoración emitida como fruto de la evaluación, sino que además conlleva un proceso permanente de seguimiento y control de la intervención. Puede, por tanto, entenderse que, en su perspectiva, el diagnóstico incluye a la evaluación.

Por otro lado, PÉREZ JUSTE (1989) defiende que el diagnóstico sirve sólo a la función de conocimiento y descripción, por lo que es la fase inicial del proceso evaluador y éste incluye a aquél. Algo más tarde, este mismo autor (PÉREZ JUSTE, 1994) matiza esta opinión señalando que las únicas diferencias que se aprecian con nitidez en los conceptos de diagnóstico y evaluación son las relativas al *profesional* que las realiza (el profesor, por un lado, que se encarga de evaluar, y el especialista, que se encarga de diagnosticar), y a la *profundidad o especialidad* de la tarea como consecuencia de esta especialización. Asimismo, también la *intención* puede diferir en ambas actividades, si bien no en exclusividad, sino en preponderancia: en el *diagnóstico*, tal intención es más

de indagación, mientras que en la *evaluación* predomina la valoración. No obstante, indagación y valoración sirven al fin de diseñar intervenciones pedagógicas y apreciar los aspectos que conducen a resultados satisfactorios o mejorables.

En nuestra opinión, la cuestión de las semejanzas y diferencias entre diagnóstico y evaluación no está zanjada en la actualidad, quizá como consecuencia de que estamos ante una disciplina que, tras años de adaptar los modelos y enfoques psicológicos, sólo recientemente se ha incorporado a la actividad educativa y está, por tanto, en plena fase de expansión.

A nuestro juicio, existen muchas semejanzas entre el diagnóstico y la evaluación, algunas de ellas especialmente relevantes para el contenido de esta obra (técnicas e instrumentos de diagnóstico y evaluación), por lo que optamos por defender la similitud frente a la diferencia. Así, siguiendo a MARTÍNEZ GONZÁLEZ (1994), a la hora de definir ambos procesos se hace referencia a la *recogida sistemática de información,* a través de técnicas que garanticen una cierta objetividad y precisión en la misma. Asimismo, la *interpretación y valoración de la información recogida con la finalidad de facilitar la toma de decisiones* son actividades que aparecen en las definiciones de ambos conceptos.

Esto implica que, tanto en el diagnóstico como en la evaluación, se acude a las mismas técnicas y procedimientos de recogida de información y, a su vez, el uso de estas técnicas se hace dentro de un proceso sistemático que debe ofrecer garantías de cientificidad. En ambos procesos, es igualmente necesario analizar e interpretar la información recogida.

Es evidente también que ni el diagnóstico ni la evaluación se realizan en el vacío, sin fines que orienten el proceso a seguir. Muy al contrario, existe siempre una finalidad (un *para*), que no es meramente producir información *per se*, sino producir información *para* decidir las intervenciones a seguir, o producir información *para* detectar problemas y corregirlos, o producir información *para* prevenir la aparición de efectos no deseados, o producir información *para* mejorar las intervenciones desarrolladas... En definitiva, tanto *diagnóstico* como *evaluación* se orientan a un *para*, es decir, se orientan a la toma de decisiones.

Ahora bien, uno de los matices que cabe establecer entre los términos diagnóstico y evaluación se refieren precisamente a la finalidad de ambas actividades. Estamos de acuerdo con MARTÍNEZ GONZÁLEZ (1994), quien plantea que el diagnóstico está enfocado siempre a la mejora, hacia el perfeccionamiento de su objeto de estudio (ya sea un producto o un proceso educativo); su misión es «proponer sugerencias y pautas perfectivas» (LÁZARO, 1986:81). Es por ello que comparte la finalidad de la evaluación formativa, pero se diferencia de la evaluación sumativa porque no adquiere una finalidad o carácter sancionador.

Como plantea MARTÍNEZ GONZÁLEZ (1994:629), «diagnóstico y evaluación forman parte, indefectiblemente, de un mismo proceso, porque no puede haber diagnóstico sin

valoración, es decir, sin evaluación». En opinión de esta autora, el proceso de diagnóstico desemboca en una interpretación de la realidad estudiada, pero esta interpretación ha de estar basada en una valoración previa de la misma, que surge, a su vez, del grado de ajuste existente entre las características de dicha realidad y las de un criterio o patrón que se toma como punto de comparación.

Por otra parte, en esta introducción queremos también hacer alguna referencia al concepto de «intervención» desde el que partimos. Como se ha dejado entrever en párrafos anteriores, entendemos que el proceso de diagnóstico y el de intervención son dos procesos altamente interrelacionados que se alimentan mutuamente y que, además, se dirigen a la educación integral del ser humano. Por otro lado, en el marco de este libro partimos de un concepto amplio de intervención, en tanto que sus contenidos no deben restringirse a la educación reglada, sino que deben abarcar también otros espacios educativos formales, no formales e informales. Así, pese a que en los párrafos anteriores nos hemos referido a los alumnos, a los profesores, a la escuela, etc., el uso de estas expresiones estaba condicionado por el hecho de ser el entorno escolar aquel en el que más incidencia ha tenido tradicionalmente el diagnóstico pedagógico. Esta disciplina se ha preocupado preferentemente por este ámbito de intervención y hemos de reconocer que sigue siendo uno de los campos más importantes en el ejercicio profesional de la Pedagogía.

Sin embargo, cada vez parece más claro que existen nuevas salidas laborales para los profesionales de la pedagogía, y numerosos ámbitos en los que este profesional desarrolla su labor (orientación profesional de personas adultas, teleformación, asociaciones y organizaciones no formales, etc.). Por esta razón, en este texto evitaremos restringir el ámbito de la pedagogía a un contexto concreto de desarrollo profesional. Por ello, procuraremos utilizar un lenguaje coherente con esta opción y rehuiremos términos como «alumnos» y «profesores», para cambiarlos por otros que den cuenta de las múltiples posibilidades de intervención de la Pedagogía, en general, y del Diagnóstico, en particular.

2. OBJETIVOS, FUNCIONES Y ÁMBITOS DEL DIAGNÓSTICO EN EDUCACIÓN

En el apartado anterior se ha hecho una breve referencia a la evolución histórica del Diagnóstico Pedagógico, poniendo el acento en las influencias que esta disciplina ha recibido de la medicina y la psicología. No obstante, el debate sobre el Diagnóstico Pedagógico en nuestro país data de mediados de los años 80, con la aparición de los textos de ÁLVAREZ ROJO (1984) y BUISÁN y MARÍN (1984). El análisis de estas y otras referencias (BRUECKNER y BOND, 1971; MARTÍ QUIRÓS, 1982) nos remiten a un diagnóstico centrado de forma casi exclusiva en el entorno escolar, la mayor parte de las veces dirigido a los sujetos y, generalmente, centrado en los trastornos y dificultades en el aprendizaje. Además de todo ello, el Diagnóstico Pedagógico ha tomado los mismos modelos y procesos utilizados en el diagnóstico psicológico para explicar el comportamiento en las situaciones de aprendizaje.

Analizando algunas de las definiciones que autores de nuestro contexto han realizado en torno al término *Diagnóstico Pedagógico*, podemos observar la evolución de este concepto desde un enfoque más correctivo a otro más preventivo y de desarrollo. Para ello, presentamos aquí algunas de estas definiciones:

> «(...) el proceso general de identificación y valoración de las necesidades de comportamiento, aptitudes, actitudes o atributos personales de un individuo dentro de un contexto dado. También puede referirse el diagnóstico a un grupo de individuos o a una situación determinada.
>
> Limitándonos al diagnóstico pedagógico, las características objeto de valoración e identificación son todos los rasgos individuales y sociales que influyen de alguna manera en el desarrollo educativo de un individuo o grupo» (ORDEN HOZ, 1969: 147).

> «1. Comprobación o apreciación del progreso del alumno hacia las metas educativas establecidas.
>
> 2. Identificación de los factores en la situación de enseñanza-aprendizaje que puedan interferir en el óptimo desarrollo individual de los escolares.
>
> 3. Adaptación de los aspectos de la situación de enseñanza-aprendizaje a las necesidades y características del discente en orden a asegurar su desarrollo continuado» (BRUECKNER y BONDER, 1971:11).

> «Proceso que, mediante la aplicación de unas técnicas específicas, permite llegar a un conocimiento más preciso del educando y orientar mejor las actividades de enseñanza-aprendizaje» (GIL FERNÁNDEZ, 1983:400).

> «El diagnóstico pedagógico es una de las actuaciones educativas indispensables para el tratamiento de los problemas que un alumno puede experimentar en un centro docente, puesto que tiene por finalidad detectar cuáles son las causas de los trastornos escolares como el bajo rendimiento académico, las conductas agresivas o inadaptadas, las perturbaciones del aprendizaje (dislexias, discalculias, etc.) y elaborar planes de pedagogía correctiva para su recuperación» (ÁLVAREZ ROJO, 1984:13).

> «Conjunto de técnicas y actividades de medición e interpretación cuya finalidad es conocer el estado de desarrollo del alumno. Nos facilita la identificación de los retrasos en el progreso de los alumnos y de sus factores causales, tanto individual como grupalmente. Es saber cómo se desarrolla, evoluciona el proceso de aprendizaje de los niños y su maduración personal en el medio escolar y familiar» (BUISÁN y MARÍN, 1984:12).

> «(...) conjunto de indagaciones sistemáticas utilizadas para conocer un hecho educativo con la intención de proponer sugerencias y pautas perfectivas» (LÁZARO, 1986:81).

«(...) trata de describir, clasificar, predecir y, en su caso, explicar el comportamiento del sujeto dentro del marco escolar. Incluye un conjunto de actividades de medición y evaluación de un sujeto (o grupo de sujetos) o de una institución con el fin de dar una orientación» (BUISÁN y MARÍN, 1987:13).

Esta evolución en las definiciones del Diagnóstico Pedagógico se hace extensible también a sus objetivos y funciones. Diversos autores (ÁLVAREZ ROJO, 1984; BUISÁN y MARÍN, 1987; GRANADOS, 1997; BUISÁN y MARÍN, 1984; MARTÍNEZ GONZÁLEZ, 1993 y PÉREZ JUSTE, 1989) han señalado los fines y propósitos a los que sirve el diagnóstico, ampliando, con el paso de los años, la lista de objetivos para avanzar desde una visión inicial de esta actividad con una función meramente correctiva (identificación y tratamiento de dificultades de aprendizaje), hasta otras funciones como la preventiva y la de desarrollo. En un intento por resumir las diversas aportaciones relativas a los objetivos del Diagnóstico Pedagógico, podemos decir que dichos objetivos son:

— *Apreciación*, o atender eficazmente los problemas surgidos como consecuencia de la variabilidad de capacidad y rendimiento de los sujetos.

— *Pronóstico*, basado en la función preventiva del diagnóstico, se dirige a que el sujeto desarrolle el máximo de sus potencialidades.

— *Prevención,* en tanto que el diagnóstico está encaminado a anticiparse a efectos futuros y a elegir correctamente entre alternativas a partir de la realidad actual.

— *Corrección* o *modificación*. El diagnóstico debe informar sobre las causas y factores condicionantes de una situación a fin de ofrecer el tratamiento de los problemas y necesidades detectados.

— *Clasificación*, en tanto que el diagnóstico se orienta a facilitar la posición relativa del sujeto en relación a otros, y su finalidad consiste en adecuar las instituciones, programas, etc. a las características diferenciales de los sujetos.

— *Reestructuración* o *reorganización* de la situación -actual o futura- para lograr un desarrollo más adecuado.

— *Comprobación* del progreso en los aprendizajes.

Asimismo, además de los objetivos a los que se orienta el Diagnóstico Pedagógico, cabría hablar de las funciones del mismo. En determinadas ocasiones, se confunden los objetivos y funciones del diagnóstico, ya que ambos presentan una dimensión teleológica (orientada a los fines y metas). No obstante, cuando hablamos de las funciones del Diagnóstico Pedagógico, en nuestra opinión, nos referimos a la finalidad o sentido último de esta actividad. Por tanto, ¿para qué sirve el diagnóstico? Así, aunque esta actividad persigue objetivos como la prevención, el desarrollo, la corrección, la clasificación, la comprobación, etc., todos estos objetivos se sintetizan en su finalidad última, ya enunciada con anterioridad, de *toma de decisiones* (PARRA, 1996). Esta toma de decisiones se lleva a cabo mediante la comparación de la situación analizada con el marco de referencia.

Por otra parte, en las definiciones presentadas anteriormente se hace referencia implícitamente a las *variables o ámbitos objeto de diagnóstico*. Es evidente que la consideración de unas u otras variables dependerá de la adscripción previa a un enfoque de diagnóstico determinado. Así, si se parte de una concepción del diagnóstico cuya finalidad sea detectar y corregir dificultades de aprendizaje, sólo nos interesará analizar ciertas variables referidas al sujeto y no otras referidas al entorno educativo y social en el que éste se ubica. Por esta razón, como plantea LÁZARO (1997), delimitar las variables del diagnóstico no es una tarea fácil, ya que se pueden adoptar diferentes criterios de selección. Existen, pues, diferentes clasificaciones que pretenden agrupar y sistematizar los ámbitos y variables del diagnóstico en educación. A continuación, vamos a presentar algunas de ellas.

BUISÁN y MARÍN (1984) aportan una clasificación de los elementos que intervienen en un diagnóstico en la cual se consideran cinco dimensiones:

— *Personal,* en la que las autoras diferencian tres ámbitos: físico/sensorial, intelectual y neurológico, y de personalidad, actitudes e intereses. Cada uno de estos ámbitos da lugar a múltiples áreas concretas que no nos detendremos a analizar.

— *Relación social*, dimensión en la que se incluye el ámbito de adaptación (personal, familiar, social, etc.).

— *Familia*, que hace referencia a la estructura y organización familiar, relaciones, formación e intereses de los padres, etc.

— *Escuela*, dimensión en la que se incluyen las condiciones y organización escolar, el clima social y las relaciones, así como la programación y la metodología.

— *Social*, que se refiere a la estructura del barrio, la relación con la comunidad, etc.

También MARTÍNEZ GONZÁLEZ (1993) plantea una clasificación de las variables del diagnóstico, agrupándolas en tres dimensiones:

— *Personal,* que incluye, a su vez, los ámbitos: biológico, psicomotor, cognoscitivo-intelectual, cognitivo, motivacional, afectivo y social. Dentro de cada uno de ellos, la autora describe múltiples áreas de exploración.

— *Académica*, que incluye los ámbitos de alumno (conocimientos, motivación, atención, etc.), profesor (formación, estilo, expectativas, etc.), aula (clima, organización, etc.), programas y medios educativos (objetivos, contenidos, recursos, evaluación, etc.) e institución escolar (aspectos físicos y organizativos, recursos, relaciones con la comunidad, etc.).

— *Socioambiental*, en la que agrupa a la dimensión familia, grupo de pares y barrio-comunidad.

En nuestra opinión, el diagnóstico en educación debe ocuparse de todas las variables relevantes en una situación educativa o, utilizando las palabras de LÁZARO (1986:88), «el

diagnóstico debe ocuparse de todo el "entramado" educativo, bien se refiera a productos, procesos, alumnos o instituciones». Por tanto, nos mostramos de acuerdo con GRANADOS (1997:45), quien describe el objeto del Diagnóstico Pedagógico como

> «cualquier cuestión, aspecto, problema, deficiencia, potencialidad y condicionamientos didácticos, organizativos y ambientales, además de las características personales y de conducta en relación con el pasado, presente o futuro de cualquier sujeto que intervenga en el proceso de enseñanza-aprendizaje».

Ahora bien, las clasificaciones anteriormente presentadas (BUISÁN y MARÍN, 1984 y MARTÍNEZ GONZÁLEZ, 1993) se han centrado en un concepto de diagnóstico estrictamente referido a los distintos niveles del sistema educativo reglado, no siendo directamente aplicables a otros campos de intervención del profesional de la pedagogía. Por esta razón, basándonos en ellas, proponemos una nueva clasificación que, sin pretender ser exhaustiva, puede aplicarse a otros contextos educativos reglados o no. Esta clasificación engloba las siguientes dimensiones:

— *Personal*, en la que se incluyen variables relativas a los sujetos de cualquier edad considerados tanto de forma individual como grupal. Dentro de ella, cabría diferenciar los ámbitos propuestos por MARTÍNEZ GONZÁLEZ (1993): biológico, psicomotor, cognoscitivo-intelectual, cognitivo, motivacional, afectivo y social.

— *Contexto social y relacional*, en la que distinguimos dos ámbitos: el contexto social inmediato (características socio-culturales y relacionales de la familia, amistades, grupo de iguales o grupo clase) y el contexto de la comunidad (características socio-culturales y relacionales del barrio, zona o comunidad del sujeto —o grupo de sujetos, programa o institución educativa—).

— *Contexto educativo*, referido al marco educativo en el que tienen lugar los procesos de diagnóstico. Dentro de esta dimensión, cabría diferenciar entre las instituciones educativas, los programas o planes de intervención, los agentes educativos, los recursos y materiales pedagógicos y didácticos, etc.

En nuestra opinión, esta propuesta recoge las principales dimensiones y ámbitos explicitados en las clasificaciones anteriormente presentadas, teniendo la peculiaridad de poder ser aplicada a otros ámbitos educativos distintos de la educación reglada.

3. ENFOQUES DE DIAGNÓSTICO

Hablar de los enfoques de diagnóstico supone hablar de diferentes tradiciones que tienen su origen fundamentalmente en el campo de la Psicología y que son las ideas que han imperado en determinados momentos históricos. Sería necesario un tema completo dedicado a esta cuestión, sin embargo, debido a la complejidad que tiene, aquí sólo queremos ofrecer una panorámica general que nos permita caracterizar diferentes procesos diagnósticos y encuadrarlos en su correspondiente tendencia.

Dado que cada modelo diagnóstico participa de una concepción teórica, de un objeto de evaluación, de unos métodos y técnicas... es esperable que la adscripción de un profesional del diagnóstico a uno de ellos condicione su actuación en cuanto al tipo de información que busca, los datos que recoge, los resultados que obtiene... En este sentido y como plantea MARTÍNEZ GONZÁLEZ (1993), el marco teórico de referencia condiciona el tipo de diagnóstico que se realiza.

La mayoría de las clasificaciones de modelos de diagnóstico (ANAYA NIETO, 1994; MAGANTO MATEO, 1996; MARTÍNEZ GONZÁLEZ, 1993; etc.) coinciden básicamente con la realizada por FERNÁNDEZ BALLESTEROS (1986) y SILVA (1982) que identifican cinco modelos (atributo, dinámico, médico, conductual radical y conductual-cognitivo) que se agrupan en dos categorías:

— La *perspectiva tradicional*, dentro de la cual se integran los modelos de atributo, dinámico y médico, que tienden a explicar el comportamiento con base en variables endógenas.

— La *perspectiva conductal*, en la que se ubican los modelos que prestan mayor atención al ambiente (conductual-radical y conductual-cognitivo).

A esta clasificación deberíamos añadir otras alternativas que han ido adquiriendo recientemente mayor relevancia y que podemos agrupar genéricamente dentro de la perspectiva cognitiva (evaluación del potencial de aprendizaje, de los estilos cognitivos, de la interacción entre iguales, etc.).

Como ya se mencionaba al principio de este capítulo, el diagnóstico es una actividad que ha estado vinculada históricamente a la Psicología. Como consecuencia de ello, el Diagnóstico Pedagógico ha estado muy influido por las distintas corrientes psicológicas dominantes: enfoque de las diferencias individuales, conductismo, psicología cognitiva. A continuación, vamos a resumir las principales características de tres formas de entender el diagnóstico, ligada cada una de ellas a las corrientes mencionadas.

3.1. Modelos tradicionales

Estos enfoques cuentan con una larga tradición y puede decirse que, si bien desde el punto de vista teórico están superados, todavía son muchos los profesionales que realizan procesos diagnósticos de acuerdo con las premisas que ahora vamos a ver. Los diferentes autores incluyen tres modelos dentro de esta perspectiva tradicional: el de atributo, el dinámico y el médico. En general, desde esta perspectiva se entiende que el comportamiento se da básicamente en función de las características organísmicas, ya sean consideradas, según cada modelo concreto, como variables biológicas, intrapsíquicas o como construcciones estructurales de la personalidad.

Sin entrar en detallar las características de cada uno de ellos, podemos sintetizar sus notas esenciales en:

— *Modelo de atributo*: Desde él se entiende que el comportamiento humano se produce en función de unos atributos intrapsíquicos que son evaluados por medio de tests y cuestionarios estandarizados.

— *Modelo dinámico:* Basado en la teoría psicoanalítica, persigue analizar la vida mental inconsciente de cada sujeto para explicar el comportamiento. La evaluación se centra en la aplicación de técnicas proyectivas.

— *Modelo médico*: Parte de la consideración de la base biológica o psíquica de los trastornos o disfunciones de la conducta humana. También basado en la aplicación de tests (en este caso, incluyendo aquellos que permiten explorar determinados procesos neuropsicológicos y sensoriales), contribuyó a la generación de entidades nosológicas consideradas como «enfermedades mentales» que requieren un tratamiento o corrección.

Dejando a un lado el modelo dinámico, que se aparta en algunos aspectos de los otros dos, las características más notorias que se pueden atribuir a la perspectiva tradicional (ANAYA NIETO, 1994; BATANAZ, 1996; BUISÁN y MARÍN, 1986; GARCÍA VIDAL y GONZÁLEZ MANJÓN, 1992; etc.) son las siguientes:

a) La unidad de análisis es el organismo, sus *rasgos o constructos internos*; estos rasgosson características psicológicas (aptitudes, personalidad, intereses...) que no se pueden observar de forma directa, por lo que son inferidos a través de las conductas observables que se consideran manifestaciones indirectas de los mismos (*signos*).

b) La evaluación y el diagnóstico se entienden como la *aplicación de tests* e instrumentos estandarizados para la medida de ciertos rasgos. Es decir, para que pueda hablarse de un proceso de diagnóstico, éste debe entenderse como una actividad relacionada con la aplicación de instrumentos de medida (diagnóstico como testación).

c) Por otra parte, la aplicación de un test se concibe como un *experimento psicológico.* Es decir, la persona que examina debe velar por que se cumplan las condiciones en las cuales el test está validado, respetando de forma escrupulosa dichas condiciones.

En caso contrario, no podemos estar seguros de que la medida que finalmente obtengamos sea buena.

d) La evaluación se orienta a *describir, clasificar y predecir* el comportamiento. En este sentido, sobre la base de las puntuaciones en diferentes pruebas, se pueden establecer pronósticos del nivel de desempeño que una persona puede alcanzar en el ejercicio de una profesión, en el marco escolar, etc.

e) Por otra parte, las unidades de evaluación son *molares*, es decir, unidades amplias y complejas (la aptitud verbal, numérica, la introversión...).

f) La evaluación se basa en el *método hipotético-deductivo*, usándose de forma preferente la metodología correlacional, a través de la cual se intentan conocer las diferencias individuales en los rasgos de la conducta humana.

g) Se pone el énfasis en la *cuantificación* como base necesaria para toda predicción del comportamiento.

h) La interpretación de los resultados es de carácter *normativo.* Aquí conviene explicar que la mayoría de los tests suelen aportar unos baremos que representan los niveles alcanzados en el rasgo en cuestión por determinados grupos poblacionales. De esta forma, la puntuación directa que se obtiene tras la aplicación de un test no tiene valor por sí misma. Nos interesa en tanto que la transformamos en otro tipo de medida (puntuación centil, típica, típica transformada, cociente intelectual, etc.) que nos permite comparar el nivel de ejecución de una persona con el del grupo de población al que pertenece. De esta forma, la población es la norma respecto a la cual se compara al sujeto. No obstante, hay que aclarar que, dentro del enfoque tradicional, se han originado otras formas distintas de interpretar los resultados: bien sea con referencia a dominios (graduación de una misma tarea según su dificultad) o a criterios (valoración del grado en que un sujeto satisface unos criterios o conductas altamente significativas). En ambos casos, estamos ante dos tendencias minoritarias, puesto que lo más habitual es la comparación normativa.

i) Como se desprende de todo lo anterior, se concede una gran importancia a la *objetividad de la evaluación.* Esto se observa en que se intentan reproducir unas condiciones de «laboratorio» en toda situación evaluadora, en la cuantificación de la medida y en su interpretación de forma normativa.

Bajo este enfoque se elaboraron y aplicaron un gran número de tests (de aptitudes, intereses, conocimientos, etc., incluyendo en este grupo a los cuestionarios de tipo estandarizado que se suelen utilizar para medir la personalidad), en un intento de estandarizar, cuantificar y realizar medidas objetivas del comportamiento (válidas y fiables).

Es de destacar, asimismo, la separación existente entre las tareas de evaluación y las de tratamiento o intervención. En los enfoques tradicionales, la evaluación termina con la comunicación de la información al sujeto (informe diagnóstico), ocurriendo en la

mayoría de los casos que la persona que realiza la evaluación no coincide con la que lleva a cabo la posible intervención.

Entre las ventajas que se señalan de la perspectiva tradicional está el hecho de haber desarrollado un gran esfuerzo por elaborar un material tipificado y estandarizado, lo que conlleva la facilidad de uso de sus técnicas y un importante ahorro de tiempo y esfuerzo, ya que se puede aplicar de forma colectiva y favorecer así la realización de diagnósticos grupales.

Sin embargo, esta perspectiva ha recibido considerables críticas, especialmente dirigidas a su falta de atención a los aspectos ambientales y su influencia en el desarrollo de la conducta, sin olvidar la falta de fiabilidad y validez de muchas de las pruebas usadas. Por otra parte, hoy día existen dudas respecto a las posibilidades de predecir eficazmente el comportamiento, objetivo éste importante para la perspectiva tradicional.

3.2. Modelos conductuales

En la medida en que el paradigma conductista comienza a ejercer su influencia en la Psicología, se producen numerosas críticas a la perspectiva tradicional y comienza a plantearse una forma diferente de entender el diagnóstico que podemos incluir bajo la denominación genérica de «evaluación conductual». En su planteamiento inicial, este enfoque trata de cambiar la tarea del evaluador desde la descripción de lo que el sujeto **es** y lo que el sujeto **tiene** (perspectiva tradicional), hacia lo que el sujeto **hace** y las condiciones bajo las cuales lo hace (SILVA, 1989).

De cualquier forma, dentro de los modelos conductuales ha habido una evolución desde el *modelo conductual-radical*, que adopta una perspectiva exógena del comportamiento al que explica sólo en virtud de las variables ambientales, hacia un *modelo conductual-cognitivo* en el que se admite la existencia de respuestas no directamente observables (cognitivas o emocionales), denominadas variables mediacionales, lo que implica que el comportamiento es el producto de la interacción persona-ambiente. Este último modelo conductual-cognitivo ha sido ampliado con las aportaciones provenientes de Bandura y de la Teoría del Aprendizaje Social, configurando lo que algunos autores (FERNÁNDEZ BALLESTEROS, 1986; BUISÁN y MARÍN, 1986; etc.) denominan como *modelo conductual-cognitivo-social*, en el que se contempla la influencia recíproca entre el organismo, el ambiente y el comportamiento, otorgando al sujeto un papel activo ante las influencias del ambiente.

Las características generales de los modelos conductuales son (BATANAZ, 1996; BUISÁN y MARÍN, 1986; GARCÍA VIDAL y GONZÁLEZ MANJÓN, 1992; etc.):

a) La evaluación de la conducta debe sustentarse en el *paradigma A-B-C*. Este paradigma expresa la idea de que la conducta (*behavior*) depende estrechamente de las variables situacionales: los antecedentes y consecuentes.

b) La conducta ha de ser evaluada en sus *contextos naturales*, a fin de que puedan conocerse esos determinantes situacionales. En este sentido, se propone una estrategia MULTIMÉTODO, MULTISITUACIÓN, MULTICONDUCTA. Como veremos a continuación, hemos de observar muchas conductas en muchas situaciones y a través de varios procedimientos.

c) El repertorio de comportamientos que es objeto de evaluación no es un signo a partir del cual se infiere un rasgo, sino una *muestra* del repertorio conductual total.

d) La conducta ha de ser evaluada *directamente* (a través de conductas concretas) e interpretarse con base en *criterios*. Esto implica que hemos de tender a hacer evaluaciones individuales e interpretar de forma individual las conductas en función de una «línea base» desde la que se establece una evaluación referida a criterios.

e) Se defiende una mayor relación entre *evaluación e intervención*. La evaluación es idiográfica y se utilizan diseños intrasujetos (N=1). La observación de las conductas en ambientes naturales deriva en la planificación de un tratamiento o intervención individual que permite el cambio conductual.

f) Por otra parte, se pone énfasis en la *explicación de variables* y en la *causalidad ambiental*. Como ya se ha mencionado, la persona realiza determinados comportamientos en función de las variables ambientales (antecedentes) y las consecuencias que dichos comportamientos le pueden reportar (consecuentes). Esto es tenido en cuenta tanto en la fase de evaluación como en la de tratamiento/intervención.

g) La evaluación se centra en conductas directamente *observables*. Se emplean las respuestas motoras, psicofisiológicas y cognitivas (aunque en el modelo conductual-radical sólo se emplean las dos primeras).

h) Por último, la *observación* es la principal estrategia de recogida de información, ya que permite un nivel de inferencia mínimo o medio. Generalmente, se suelen utilizar observaciones directas y sistematizadas en las que se parte de una definición previa de las conductas (categorías) a observar. Sin embargo, también se pueden utilizar otras estrategias observacionales como los autorregistros y los autoinformes y, en algunos casos, la entrevista.

En definitiva, el enfoque de la evaluación conductual presenta grandes diferencias respecto a la perspectiva tradicional. Estas diferencias se refieren a la forma de entender el objeto de evaluación, las fases del proceso de diagnóstico, los instrumentos, el tratamiento posterior, etc.

Las principales críticas a esta perspectiva se centran en el reduccionismo derivado de eliminar los procesos internos como unidad de estudio —crítica que se debe aplicar en mayor medida al modelo conductual radical—. Ahora bien, una de las grandes aportaciones de este enfoque es, sin duda, la estrecha vinculación que se defiende entre la evaluación y la intervención, las cuales forman parte de un continuo inseparable.

3.3. Modelos derivados de la psicología cognitiva

El auge de la psicología cognitiva ha originado una nueva aproximación a la cuestión del diagnóstico y la evaluación, si bien, muchas de las aportaciones del cognitivismo han sido asumidas por la perspectiva conductual integrándose en el modelo conductual-cognitivo y en el conductual-cognitivo-social. Al estar todos estos modelos vinculados con planteamientos cognitivistas, nosotros, al igual que autores como BATANAZ (1996), IBARRA (1999), MAGANTO (1995) y MAGANTO MATEO (1996), los agruparemos en una única perspectiva denominada *cognitivista.*

No obstante, dentro de la perspectiva cognitivista, no existe consenso entre los distintos autores revisados en torno a las principales tendencias a incluir en este grupo. En términos generales, se señalan las siguientes contribuciones cognitivistas al diagnóstico: la pedagogía operatoria, la evaluación del potencial de aprendizaje, los estilos cognitivos, la interacción entre iguales y la evaluación del clima de clase.

El denominador común de los modelos cognitivistas es la consideración de la persona como un ser activo que busca, selecciona, codifica y almacena la información y planifica y toma decisiones. La perspectiva cognitiva pone el énfasis en conocer cómo la persona reacciona ante una situación, cobrando gran relevancia los aspectos afectivos y actitudinales.

Revisar todas estas tendencias escapa a nuestra intención, por lo que nos limitaremos a destacar brevemente los rasgos más significativos de las que consideramos más relevantes: la evaluación del potencial de aprendizaje y de los estilos cognitivos.

En la **evaluación del potencial de aprendizaje,** basada en la teoría potencial de Vygotski, destacan las aportaciones de R. Feuerstein. El objetivo central de este enfoque es evaluar el nivel de educabilidad —modificabilidad cognitiva— del sujeto, informando de en qué medida dicho sujeto se beneficiaría de un programa de enriquecimiento cognitivo. Este enfoque, si bien se opone al uso de tests en su aplicación más ortodoxa, basa la evaluación en el uso de estos instrumentos, ahora bien, cambiando considerablemente las condiciones de aplicación y de interpretación de resultados. Por otra parte, aunque se opone a la conceptualización conductista de que sólo interesan las conductas directamente observables, retoma de este enfoque su principio de que es preciso fundamentar los tratamientos correctivos en la propia evaluación.

Veamos con mayor detalle algunas de las características principales de esta forma de entender el diagnóstico y la evaluación (GARCÍA VIDAL y GONZÁLEZ MANJÓN, 1992; MAGANTO MATEO, 1996; BUISÁN y MARÍN, 1986; etc.):

— La evaluación se centra en los *procesos mentales y cognitivos* y no tanto en los productos (principal preocupación de los defensores del enfoque tradicional). Desde este punto de vista, y a diferencia del conductual, el interés reside en los procesos

de pensamiento humano, los cuales son de carácter interno y, por lo tanto, no pueden observarse directamente, aunque podemos inferirlos a través de la evaluación.

— El objetivo del diagnóstico es evaluar la capacidad de aprendizaje, para lo cual se sigue una *metodología experimental intrasujeto* que sigue un esquema de test-entrenamiento-retest, a partir de la cual se persigue conocer la ganancia del sujeto debida al tratamiento.

— Para poder inferir la actividad cognitiva del ser humano es necesario poner en práctica un proceso de evaluación basado en la *mediación*. Esto implica algunos cambios en la forma de llevar a la práctica el diagnóstico respecto a enfoques anteriores:

- Cambios en la *relación examinador/examinado*. Frente a la situación de evaluación clásica, propia del modelo de atributo, en la que el examinador procura adoptar un papel neutro para poder reproducir las condiciones experimentales de aplicación de un determinado test, el examinador desde esta perspectiva demuestra activamente su interés en el trabajo del sujeto. Su rol consiste en favorecer una situación de interacción en la que se facilite la *mediación*; es decir, promover cierto tipo de interacciones que nos permiten conocer cuáles son los procedimientos mentales que el sujeto pone en juego para resolver una tarea determinada.
- Cambios en la *estructura/situación del examen*. En el sentido de que la situación de «prueba» o «experimento» pasa a ser una situación de interacción que nos permite promover los procesos cognitivos internos.
- Cambios en la *forma de interpretar los resultados*. No se busca la comparación con la norma. Los errores del sujeto nos aportan información, en tanto que nos permiten comprender su funcionamiento cognitivo. Las puntuaciones se calculan a través de procedimientos distintos que vienen a informarnos del grado de modificabilidad cognitiva de la persona (grado en que ésta se puede beneficiar de un determinado tratamiento educativo).

Puesto que la evaluación se centra en procesos cognitivos que no son directamente observables, es una tarea bastante compleja que requiere un alto nivel de formación para su correcta realización. Así, aunque *los tests y la observación* son las principales estrategias de recogida de información, éstas tienen sentido si se facilita la mediación y la actividad cognitiva del sujeto, por lo que es necesario saber cuáles son los mecanismos para que así sea. En este sentido, y como ya hemos mencionado, se trata de una forma muy heterodoxa de aplicar los tests.

Algunas de las estrategias de evaluación del potencial de aprendizaje con las que contamos en la actualidad son el LPAD (*Learning Potential Assessment Device*), elaborado por el propio Feurstein, y el EPA (Test de Evaluación del Potencial de

Aprendizaje), realizado bajo la dirección de Rocío Fernández Ballesteros. Aunque en ambos instrumentos el procedimiento de evaluación es muy diferente, sin embargo, esta actividad se realiza en los dos con las características generales antes descritas.

Otra tendencia importante dentro de la perspectiva cognitivista de la evaluación es la de **estilos cognitivos**. Este enfoque también se sustenta en un concepto de evaluación de los procesos más que de los productos. El interés reside en conocer cómo actúa el sujeto desde una perspectiva holística que no sólo tiene en consideración los aspectos propiamente cognitivos (percepción, razonamiento, resolución de problemas), sino también los referidos a la personalidad.

Entre los rasgos más destacados de esta aportación, podemos señalar los siguientes (BUISÁN y MARÍN, 1986; IBARRA, 1999; MARTÍNEZ GONZÁLEZ, 1993; SOBRADO y OCAMPO, 1997):

— Se entienden los estilos cognitivos como *hábitos de procesamiento de la información* o modos característicos del sujeto de operar que poseen una cierta constancia a través de diferentes áreas de contenido (reflexividad-impulsividad, dependencia-independencia, convergencia-divergencia, etc.).

— La evaluación de estos estilos se enfoca hacia la *determinación cualitativa de diferencias individuales.* A través de metodologías de comparación interindividual se pretende hacer aflorar las diferencias en el procesamiento ante tareas relevantes (percibir, pensar, resolver problemas, relacionarse con los demás, aprender, etc.).

— Como recursos evaluadores se utilizan *pruebas* como la de figuras enmascaradas (dependencia-independencia de campo), tests de búsqueda de semejanzas y diferencias (reflexividad-impulsividad), pruebas de tareas que permitan la creatividad (pensamiento convergente-divergente), etc. A menudo, también se incorpora la observación directa de las realizaciones del sujeto.

La evaluación de los estilos cognitivos tiene una gran aplicación en el marco educativo, ya que nos permite conocer cómo cada persona concreta percibe y procesa la información, siendo ésta una información muy útil para orientar los programas educativos.

Aunque no vamos a detenernos en ellas, son también muy interesantes para el campo educativo las tendencias *de interacción entre iguales* y la *percepción del clima de clase*. En la interacción entre iguales se persigue analizar las interacciones que se producen en el alumnado cuando realiza ciertas tareas; es decir, conocer cómo se ejecuta la tarea en la interacción y las situaciones de interacción que se producen. Las técnicas de observación han sido las más utilizadas en esta tendencia.

Por su parte, en la percepción del clima de clase se persigue evaluar el ambiente de clase y su influencia sobre el aprendizaje de los sujetos y grupos. Las técnicas utilizadas han sido muy diversas (observación, técnicas sociométricas y escalas o inventarios), adoptándose tanto un enfoque correlacional, que relaciona, fundamentalmente, variables

ambientales e instructivas, como una metodología interpretativa, centrada en analizar los procesos por los cuales los miembros que comparten un ambiente definen y dirigen sus actividades cotidianas (MARTÍNEZ GONZÁLEZ, 1993).

4. APORTACIONES DESDE LA INTERVENCIÓN ORIENTADORA: PRINCIPIOS Y MODELOS PARA EL DIAGNÓSTICO EDUCATIVO

Es evidente que cada profesional del diagnóstico y la evaluación, según su formación y preferencias personales, se adscribirá a uno de los enfoques presentados en el apartado anterior. Sin embargo, todos los enfoques presentados provienen del campo de la Psicología y, aunque tienen utilidad en el ámbito educativo, están pensados desde las características propias de esta disciplina y adecuados a las prácticas habituales de los psicólogos.

Sin abandonar por completo estas aportaciones, resulta, a nuestro juicio, necesario definir y caracterizar el diagnóstico desde una perspectiva más educativa, teniendo en cuenta las características de las Ciencias de la Educación y las prácticas habituales de los profesionales de la Pedagogía. Sin embargo, ésta es una tarea que excede ampliamente el objetivo de esta obra, y que necesitaría de un debate profundo entre todos aquellos que se dedican al diagnóstico en educación.

No obstante, pensamos que sí resulta factible documentar aquí algunos principios básicos hacia los que debe tender el diagnóstico pedagógico. Tendremos para ello en cuenta las aportaciones revisadas en apartados anteriores, donde hacíamos referencia a los objetivos, fines, etc. del proceso de diagnóstico. Asimismo, nos parece fundamental considerar las aportaciones realizadas desde una disciplina con la que el diagnóstico ha tenido históricamente una gran vinculación. Nos estamos refiriendo a la Orientación Educativa.

Diversos autores de nuestro contexto (BUISÁN y MARÍN, 1986; MARÍN, 1997; etc.) han resaltado en numerosas ocasiones que el diagnóstico se incardina dentro del proceso de orientación educativa, en tanto que proporciona las claves para la intervención. Hasta tal punto que forman un todo unitario, porque sin el diagnóstico la orientación no tiene rigor, y el diagnóstico sin la orientación no tiene sentido (PÉREZ JUSTE, 1990). Por lo tanto, ambos procesos se entienden como interactivos, no lineales, requiriéndose mutuamente durante una intervención, de ahí que compartan objetivos y funciones (MARTÍNEZ GONZÁLEZ, 1993).

Es por ello que pensamos que los objetivos, principios y modelos de intervención que hoy día se aceptan y defienden en la orientación educativa, deben también inspirar los procesos diagnósticos, relacionados, al fin y al cabo, con aquélla. En este apartado, y sin ánimo de ser exhaustivos en la revisión, queremos sintetizar estos objetivos y principios de la intervención orientadora actual, a fin de que sean útiles también para la realización de procesos diagnósticos.

Una de las primeras ideas que podemos considerar es la que se refiere a los **fines y objetivos** a los que la intervención orientadora debe servir. Tomando como referente lo que propone la LOGSE[1], estos fines son (MEC, 1990):

1. Contribuir a la personalización de la educación.
2. Ajustar la respuesta educativa mediante la adaptación del currículum.
3. Cooperar con los aspectos orientadores de la educación.
4. Favorecer los procesos de madurez personal, de la propia identidad, de los propios valores y la toma de decisiones.
5. Garantizar los elementos educativos más diferenciados, especializados, actualizados y progresivos.
6. Prevenir las dificultades de aprendizaje.
7. Asegurar la continuidad educativa en áreas, ciclos, etapas o centros.
8. Contribuir a las adecuadas relaciones e interacciones entre los distintos sectores de la comunidad educativa (profesorado, alumnado y familia), asumiendo la mediación y la negociación precisas.
9. Impulsar la innovación educativa.

Es conveniente prestar atención a los presupuestos que subyacen en las finalidades enunciadas. Para empezar, se defiende una *concepción global e integrada* de los sujetos y de los procesos educativos. Asimismo, se considera la orientación como un *subsistema de la educación*, y no tanto como un sistema externo e independiente del mismo. Por otra parte, se pone el acento en la *prevención* de las posibles dificultades y en el *desarrollo* de las potencialidades de los educandos.

Se destaca también una *perspectiva sistémica* de la orientación: lo individual y lo institucional como ámbitos de actuación. Es decir, la intervención orientadora no debe enfocarse en exclusividad a «cambiar al sujeto», sino a cambiar el sujeto y los contextos o sistemas con los que éste interactúa.

En realidad, el hecho de que la LOGSE adopte esta perspectiva no viene sino a reflejar el cambio de modelos y enfoques que ha tenido lugar en la Orientación como disciplina científica, desde que naciera a principios de siglo hasta nuestros días. Así, la orientación surge fuera de la escuela y se entiende como una «ayuda extra y especializada» a determinados sujetos que presentan dificultades y problemas. Hoy día, se ha producido un cambio sustancial en este enfoque y la mayoría de los autores defienden un concepto de *orientación educativa,* en el que se potencia una perspectiva según la cual la orientación no es un «añadido», sino que forma parte de la acción educativa y, como tal, se dirige a los mismos fines y sujetos.

Desde una *orientación educativa*, se entiende que la intervención debe dar cuenta de un conjunto de **principios** que, por otra parte, están integrados plenamente en el

planteamiento de la LOGSE. En esencia, estos principios son tres (RODRÍGUEZ ESPINAR et al., 1993):

— El *principio de prevención*: que supone adelantarnos a determinados eventos, bien para evitar que ocurran, bien para minimizar sus efectos. La introducción del principio de prevención en la intervención psicopedagógica implica adoptar una perspectiva de orientación proactiva, que va dirigida a toda la población, y no a un grupo determinado de alumnos y alumnas.

— El *principio de desarrollo*: la persona va estructurando su personalidad, capacidades, habilidades y destrezas en su interacción con el medio. La orientación, por tanto, debe favorecer un contexto motivador que estimule estos cambios, fomentando el desarrollo de las competencias necesarias para abordar cada una de las fases evolutivas (ÁLVAREZ, 1995). Una consecuencia de la aceptación de este principio es la consideración de la orientación como un proceso continuo dirigido al desarrollo integral de la persona (RODRÍGUEZ ESPINAR et al., 1993). Por otra parte, la dimensión educativa de la orientación se pone de manifiesto cuando los objetivos educativos se persiguen a lo largo de todo el proceso de enseñanza-aprendizaje, y no como actuaciones marginales y puntuales.

— El *principio de intervención social*: la intervención orientadora no debe tener sólo en cuenta al individuo, sino todo el contexto social en que éste se desenvuelve. Este principio tiene también importantes connotaciones en la práctica orientadora, ya que implica que no se trata sólo de que la persona se adapte al medio, sino modificar algunos aspectos del contexto que obstaculizan la realización personal y desarrollar las condiciones y oportunidades más favorables en el ambiente (ÁLVAREZ, 1995). En este sentido, la orientación se constituye en un agente de cambio y el análisis de contextos en una actividad absolutamente imprescindible a la hora de enfocar la intervención.

En opinión de muchos autores (e.g. MAGANTO MATEO, 1996; MARTÍNEZ GONZÁLEZ, 1993; SOBRADO Y OCAMPO, 1997), estos principios son también aplicables al diagnóstico y la evaluación y pueden, en consecuencia, inspirar y orientar esta actividad. Esto no excluye la necesidad de intervenir ante problemas o dificultades ya manifestadas, adquiriendo en este caso el diagnóstico un matiz correctivo o terapéutico.

GARCÍA VIDAL y GONZÁLEZ MANJÓN (1992), recogiendo en parte los ya mencionados principios de prevención, desarrollo e intervención social, hacen una propuesta en la que aparecen otros nuevos principios que deben inspirar la orientación educativa. Estos principios, por otra parte, están recogidos en el modelo de orientación educativa propuesto desde la LOGSE y, en nuestra opinión, pueden también aplicarse al diagnóstico y la intervención en otros contextos educativos distintos a la educación reglada. Los principios en cuestión son (GARCÍA VIDAL y GONZÁLEZ MANJÓN, 1992):

— El *principio de planificación*, que hace referencia a que los procesos orientadores deben estar sujetos a una planificación y desarrollarse a través de programas de

intervención. En la medida en que la actividad orientadora es planificada, es posible dar cabida a los principios de prevención, desarrollo e intervención social.

— El *principio de prevención*, en los términos ya comentados. Así, para que nuestra actuación sea preventiva ha de estar dirigida a dos objetivos básicos: la disminución de los factores de riesgo y la optimización de los procesos educativos, con lo cual GARCÍA VIDAL y GONZÁLEZ MANJÓN (1992) asumen el principio de desarrollo como parte de la prevención.

— El *principio de sistemización*, que viene a significar lo mismo que el principio de intervención social. Desde esta perspectiva, hay que considerar que los destinatarios de la intervención no son exclusivamente las «personas con problemas», sino también los «sistemas» (instituciones, centros y otras estructuras). Esto implica varias cosas: a) por una lado, considerar los «casos» como sistemas en los que inciden múltiples variables relacionadas entre sí; b) por otro, realizar «análisis de contextos» y no tanto «análisis de sujetos»; c) y por último, dar respuestas educativas a las necesidades detectadas tanto a nivel individual como institucional/grupal.

— El *principio de curricularidad*, en tanto que el currículo se debe considerar como el medio por excelencia para la intervención orientadora, de tal forma que las intervenciones deben vehicularse a través de él. Puesto que la orientación es un sistema de la educación, entre la intervención orientadora y el currículo debe existir una estrecha relación, tanto desde la perspectiva del diagnóstico como de la propia intervención.

— El *principio de cooperación*, de acuerdo con el enfoque de la orientación y del diagnóstico que venimos defendiendo, es evidente que éstos no son tarea exclusiva de un determinado profesional, sino que requieren de la intervención de todos los agentes educativos, cada uno de ellos con unas responsabilidades diferentes y actuando en niveles distintos.

Además de estos principios, puede hablarse de diferentes **modelos de intervención** orientadora. Así, dadas las relaciones existentes entre el diagnóstico y la intervención, y la vinculación tradicional existente entre las actividades diagnósticas y orientadoras, los modelos que han prevalecido en orientación han condicionado también el tipo de diagnóstico que se ha venido realizando.

Desde un punto de vista general, se han identificado dos modelos básicos de intervención en orientación (ÁLVAREZ, 1991 y 1995; RODRÍGUEZ ESPINAR et al., 1993), modelos que han implicado diferentes formas de entender y llevar a la práctica el diagnóstico educativo. Estos modelos son el denominado modelo de servicios y el de programas.

Con anterioridad a la aplicación de la LOGSE, en nuestro país prevalecía un *modelo de servicios* como estructura de apoyo a la labor educativa. Desde este enfoque, la

orientación era desarrollada por determinados equipos de profesionales subvencionados a través de organismos públicos (consejerías de educación, ayuntamientos, etc.), aunque el trabajo desarrollado por los profesionales de la orientación que trabajan dentro de gabinetes y consultas privadas también se asemeja mucho a la actuación desde el modelo de servicios.

Entre las principales características de esta forma de intervenir, podemos señalar las siguientes (ÁLVAREZ, 1991 y 1995; RODRÍGUEZ ESPINAR et al., 1993):

— Su actuación viene predeterminada por un conjunto de funciones a realizar. No se establecen objetivos de intervención y, por lo tanto, se carece de una planificación. En este sentido, se atiende a los problemas que se van manifestando, según las funciones y profesionalización de los miembros del equipo.

— La orientación es entendida como una tarea exclusiva del especialista (pedagogo, psicólogo, etc.). A él acuden los casos a los que el profesor de aula no puede dar respuesta.

— Se trabaja, por tanto, de forma descontextualizada. Es decir, no se atiende al alumnado en su contexto de referencia, sino que se actúa aisladamente sobre su «problema» o «dificultad».

— Es un enfoque fundamentalmente terapéutico y remedial. Los equipos tratan de remediar los problemas de ciertos alumnos, aquellos que manifiestan dificultades en el proceso de escolarización, mientras que el resto del alumnado no se beneficia de los servicios. Por otra parte, las respuestas orientadoras tienen un carácter correctivo: «solucionar» problemas, «tapar» agujeros.

Ante esta forma de entender la orientación, las tareas del diagnóstico consisten en delimitar de forma precisa el tipo de problema o deficiencia, de forma que se promueva una intervención que va a ser desarrollada con frecuencia fuera del aula ordinaria.

Dadas las limitaciones de este enfoque, muchos autores y autoras (ÁLVAREZ, 1991; RODRÍGUEZ ESPINAR et al. 1993; etc.) han venido insistiendo en la necesidad de cambiar el modelo de intervención para tender hacia una *orientación por programas*. Esto supone un importante cambio en la forma de concebir la orientación, ya que en el modelo de servicios subyace un concepto de orientación entendida como una actividad puntual, marginal y fundamentalmente de ayuda, mientras que el modelo de intervención por programas se orienta hacia un concepto de orientación *educativa,* e intenta responder a los principios que anteriormente mencionamos.

Recogiendo las aportaciones de la investigación educativa y psicológica, la LOGSE introduce cambios significativos en la forma de entender la orientación, adoptando una perspectiva por programas. Esta nueva perspectiva se caracteriza por (ÁLVAREZ, 1991; RODRÍGUEZ ESPINAR et al. 1993; etc.):

— La actuación orientadora se realiza en función de las necesidades existentes. Esto implica que se parte de una evaluación de necesidades a partir de la cual se determinan los objetivos de la intervención y se planifican las estrategias, medios y recursos para su consecución. Por eso, la orientación se lleva a cabo a través de programas de actuación.

— La determinación de unos objetivos y estrategias permite que pueda realizarse la evaluación de la actividad orientadora, lo cual no resulta posible en un modelo de intervención por servicios.

— La intervención va dirigida a todo el alumnado, tanto de forma individual como grupal. No se circunscribe, por tanto, al alumnado con dificultades y problemas.

— Los programas de orientación son procesos a través de los cuales se intenta responder a los principios de prevención, desarrollo e intervención social. Por eso, la actuación orientadora no se limita a resolver problemas y dificultades (enfoque remedial), sino que adopta un enfoque de carácter preventivo y de desarrollo.

— Los programas se realizan dentro de los centros e instituciones educativas. Se tiene, por tanto, en cuenta el contexto de la intervención.

— Por otra parte, los programas son estrategias complejas que requieren de una intervención a múltiples niveles. Por ello, la orientación no es la actividad de un especialista, sino una tarea compartida por todos los profesionales de la educación y agentes educativos. Cada uno tiene unas responsabilidades diferentes pero de todos es necesaria la cooperación.

Bajo estos supuestos, la forma de entender y llevar a la práctica el diagnóstico educativo ha experimentado notables cambios. Ya no repercute sólo en el alumnado, ni tampoco en un subconjunto concreto de éste; tampoco se entiende como una actividad exclusiva del especialista, sino como una tarea que se realiza en diferentes niveles y en la que distintos agentes adquieren distintas finalidades.

De todas formas, queremos destacar que, en nuestra opinión, los dos modelos de intervención orientadora aquí analizados (por servicios y por programas) son, más que modelos «puros», los extremos de un continuo, de tal manera que las distintas intervenciones concretas se acercan más a uno u otro extremo del continuo. Basta decir que, pese a la asunción de un modelo de programas tras la aplicación de la LOGSE, el sistema de orientación de nuestro país, sigue contando con determinados «servicios» (equipos de sector, departamento de orientación...) encargados de implementarlo.

Estos principios y modelos de intervención, hoy día asentados en la práctica de la orientación, deben ir inspirando cada vez más la actividad diagnóstica. Algunos autores han defendido este planteamiento, sugiriendo que el diagnóstico debe tomar estas características de la intervención orientadora. Por ejemplo, GARANTO (1997) sugiere que el diagnóstico en educación debe evolucionar hacia un planteamiento:

— *Dinámico*, en el que cobren mayor relevancia los contextos y sistemas en los que interactúan las personas.

— *Procesual*, que no se limite a momentos puntuales, sino que abarque todo el proceso educativo.

— *Integrador*, en tanto debe contemplar no sólo todas las facetas del desarrollo humano, sino también la interacción de éstas con el contexto.

En términos similares se expresa MARÍN (1997), quien plantea que estamos asistiendo a un redescubrimiento de la tarea educativa que no puede concebirse sólo como instrucción. Se observa así un interés creciente por devolver a la educación su papel de agente impulsor del desarrollo personal. A nivel del diagnóstico, este replanteamiento de la tarea educativa implica abrir su significado a lo «pedagógico», abarcando más aspectos que el proceso de enseñanza-aprendizaje. Teniendo esto en cuenta, un modelo de diagnóstico más adecuado que los que existen en la actualidad debería tener el siguiente conjunto de rasgos (MARÍN, 1997):

— Tener como objetivo prioritario ayudar al educando, haciendo énfasis en la valoración más que en la aplicación de tests.

— Tener en cuenta la validez ecológica, de forma que el diagnóstico se realice en el mismo contexto en que se desenvuelve el educando y no en situaciones extrañas.

— Es importante profundizar en el objeto de diagnóstico y en las variables que se seleccionan, no sólo en sí mismas, sino dentro de unos modelos teóricos que den respuesta a las características individuales.

— Los instrumentos utilizados han de ser sensibles a las diferencias individuales y grados de desarrollo. Se necesita, además, utilizar múltiples medidas y técnicas en los procesos de diagnóstico.

— El proceso de diagnóstico ha de formar parte del ambiente natural de aprendizaje.

— Es interesante incorporar al diagnóstico nuevos procedimientos de recogida de información como, por ejemplo, la valoración de ejecuciones.

Hace ya bastantes años, RODRÍGUEZ ESPINAR (1982) señalaba un conjunto de aspectos que deben caracterizar al diagnóstico en educación:

— Debe estar abierto a la utilización de múltiples metodologías de obtención de información.

— Debe poner de manifiesto potencialidades o limitaciones.

— Siempre debe intentar una explicación causal.

— Las medidas realizadas en su proceso deben estar referidas a una norma o criterio.

— Debe ponerse especial atención a la relevancia y adecuación de los factores a diagnosticar.

— Ha de evitarse la exclusividad de interpretación centrada únicamente en el sujeto y dar un mayor énfasis a los factores ambientales que enmarcan cada situación.

En relación con la primera característica señalada por RODRÍGUEZ ESPINAR (1982), y aunque en un apartado posterior profundizaremos más en esta cuestión, queremos destacar que una de las principales carencias que observamos en los enfoques diagnósticos revisados anteriormente es la preferencia y casi exclusividad con que se usan las técnicas de diagnóstico. Así, en la perspectiva tradicional, sólo los tests permiten hacer diagnóstico; en el enfoque conductual, sucede lo mismo con la observación en todas sus modalidades... Como veremos en sucesivos capítulos, cada técnica aporta unas ventajas y limitaciones al proceso de diagnóstico. Cuando trabajamos en educación, dada la amplitud y complejidad de las variables implicadas, la preferencia en exclusiva de una técnica puede resultar una opción reduccionista, que no contribuye a una comprensión y descripción en profundidad del objeto de diagnóstico.

Todas estas consideraciones no son sino notas que pretenden caracterizar el proceso de diagnóstico educativo. Es evidente que la lista de ideas aquí aportadas no es ni mucho menos exhaustiva. Su finalidad no es otra más que potenciar la reflexión y el debate sobre la forma en que podemos y debemos hacer diagnóstico en el campo de la educación, pasando a un segundo plano las aportaciones, sin duda valiosas, que esta disciplina ha tenido desde el campo de la Psicología.

5. FASES EN EL PROCESO DE DIAGNÓSTICO

En este apartado vamos a presentar una idea general de las diversas fases y tareas implicadas en el proceso de diagnóstico y evaluación. En realidad se trata de ofrecer una panorámica general de todo el proceso, sin profundizar en cada una de las fases que aquí vamos a mencionar. No obstante, recogemos la aportación de IBARRA (1999), quien plantea que, puesto que el diagnóstico es un proceso, no resulta conveniente señalar en él etapas o fases puntuales, lineales o inconexas. Cabe hablar mejor de un proceso general que tiene diversos momentos, necesarios y relacionados, que en ocasiones no siguen una sucesión lineal, sino circular o en espiral, según la retroalimentación del proceso así lo vaya aconsejando.

Por otra parte, hemos de tener en cuenta que cada proceso de diagnóstico puede presentar unas características diferenciales en función de aspectos como: el modelo teórico al que se adscriba el profesional que lo realiza, las variables o dimensiones en que se centre, el hecho de que el diagnóstico afecte a un sujeto, a un grupo, a un programa, a una institución, etc. Por lo tanto, lo que veremos a continuación es una propuesta general que, en cada situación concreta, requerirá las correspondientes modificaciones y adaptaciones.

Asimismo, puesto que anteriormente hemos defendido los vínculos existentes entre el diagnóstico y la intervención educativa, en nuestra propuesta de proceso de diagnóstico hemos tenido en cuenta esta relación, articulando los distintos momentos en función de la misma. Cabe aclarar, recogiendo la idea de circularidad que mencionábamos, que determinados procedimientos de recogida de información permiten combinar las actividades diagnósticas con las de intervención. Así, cuando se aplican estos procedimientos (por ejemplo, la historia de vida, el grupo de discusión, etc.), no sólo estamos recogiendo información, sino también propiciando la reflexión y la toma de conciencia respecto a la situación problemática que suscita el diagnóstico y, por tanto, comenzando la intervención.

De acuerdo con todo esto, y tomando algunas de las aportaciones realizadas por diversos autores (BUISÁN y MARÍN, 1987; GARCÍA VIDAL y GONZÁLEZ MANJÓN, 1992; MAGANTO, 1996; IBARRA, 1999; etc.), el proceso de diagnóstico educativo comprende las siguientes tareas y momentos generales:

1. Planificación y delimitación inicial

En este primer momento del proceso es necesario hacer una reflexión inicial sobre la situación que origina el diagnóstico y la forma de proceder ante la misma. Se trata de definir el objetivo u objetivos del diagnóstico, delimitar quién, quiénes o qué va a ser evaluado, cuáles van a ser las áreas a explorar, las técnicas a utilizar, las personas implicadas en la evaluación, etc.

Para responder a todos estos interrogantes será preciso en muchos casos hacer una *recopilación inicial de información* a partir de la cual, el profesional puede formarse una imagen más exhaustiva y completa de la situación, problema o necesidad que suscita el proceso de evaluación. En esta recogida de información previa es conveniente además utilizar diversas fuentes (sujeto, familia, agentes educativos...) que contribuyan a percibir las diferentes manifestaciones del problema o necesidad.

En determinados marcos educativos sucede a menudo que la demanda que llega al profesional del diagnóstico es diferente a la situación que conviene analizar y mejorar. Un ejemplo clásico de esto lo encontramos en muchos orientadores de enseñanza primaria o secundaria, que reciben múltiples peticiones de diagnóstico porque algunos alumnos «tienen problemas». Es probable que, a poco que el orientador recopile información, se percate que no se trata de «un problema del alumno», sino de necesidades educativas que surgen de la relación alumno-profesor, alumno-metodología, etc.

Es por esta razón que es necesario hacer una recopilación previa de información que permita fijar de forma más realista los objetivos y/o hipótesis de partida que van a orientar el proceso de diagnóstico. Una ventaja añadida de proceder de esta forma es suscitar la motivación e implicación en las personas afectadas, que así no se desentienden del proceso de diagnóstico, sino que se percatan que éste va a requerir su participación.

Tras la realización de las tareas correspondientes al primer momento, dispondremos de un diseño o plan de lo que va a ser el proceso a seguir.

2. Recogida de información

En esta segunda fase se trata de recoger el máximo de información posible sobre la situación objeto de estudio, guiándonos por los objetivos y/o hipótesis y el plan de trabajo formulados en el momento anterior. Este plan debe ser, no obstante, flexible para dar cabida a nuevos objetivos que pueden surgir como consecuencia de una recogida más intensiva de información.

Conviene hacer uso de diferentes técnicas y procedimientos (observación, entrevistas, cuestionarios, producciones del sujeto, tests y pruebas, análisis de documentos, etc.). Aunque dependerá del tipo de proceso de diagnóstico que estemos realizando, resulta bastante adecuado combinar estrategias tanto cualitativas como cuantitativas. Las primeras, al ser abiertas y flexibles, nos permitirán conseguir una cierta profundidad al tiempo que una aproximación a la perspectiva de las personas implicadas. Las segundas, por otra parte, nos permitirán operativizar las variables a estudiar y trabajar con grandes cantidades de información.

Asimismo, y como en el momento anterior, es importante utilizar la información procedente de distintas fuentes, lo que nos va a dar una visión más completa y extensa de la situación. Por otra parte, conviene explorar múltiples áreas o dimensiones y aunque

algunas de ellas serán objeto de un análisis más exhaustivo, otras pueden ser de gran ayuda en la explicación de la situación.

3. Interpretación y valoración de la información

Se trata ahora de analizar la información recogida en las fases anteriores. Los procedimientos de análisis pueden ser diversos, dependiendo de las estrategias de recogida de información utilizadas y los propios fines del diagnóstico. Esto implica tareas de síntesis —resumir y condensar la información recabada—, de interpretación y de valoración de la información, intentando describir y explicar la situación analizada.

De cualquier forma, interpretar la información es una actividad compleja. En ella debemos relacionar la información obtenida por distintas vías y referidas a distintas variables. Asimismo, esta información debe ser contrastada con algún referente (objetivos educativos), a fin de formular los juicios de valor y estimar las necesidades prioritarias para la intervención.

4. Toma de decisiones y propuesta de intervención

Las conclusiones del momento anterior nos han permitido delimitar un conjunto de necesidades priorizadas sobre las que hay que intervenir. Se trata ahora de elaborar un programa de intervención coherente con la valoración de la información que realizamos en la fase anterior. Puesto que disponemos de información abundante al respecto, estamos en condiciones de decidir sobre los objetivos a trabajar, las metodologías, actividades, recursos, materiales didácticos, etc. más adecuados para superar las necesidades detectadas, los agentes encargados de las distintas tareas asociadas al programa, etc.

Por otra parte, es necesario delimitar las responsabilidades de cada agente educativo respecto al programa, ya que es evidente que no será siempre el profesional del diagnóstico el encargado de aplicarlo. Por lo tanto, se trata también de decidir quién hace qué y cómo se puede facilitar la coordinación.

5. Seguimiento/evaluación

Una vez que hemos puesto en marcha el programa no podemos desentendernos de él, sino que tiene que ser objeto de revisión continua, de una evaluación de su desarrollo (proceso) y sus resultados (productos). Ésta es la única forma de comprobar la eficacia o inadecuación de la propuesta de intervención e, indirectamente, del proceso de diagnóstico que la suscita.

La realización de los correspondientes informes y la comunicación de los resultados pueden tener lugar en distintos momentos de los anteriormente expuestos. En la mayoría de los casos, estas tareas se realizan con posterioridad a la interpretación y valoración de la información, aunque consideramos recomendable hacerlo tras la toma de decisiones y la propuesta de intervención. De esta forma, el informe da cuenta, no sólo de las

necesidades detectadas, sino también de las actuaciones que se van a poner en práctica para superarlas. De cualquier forma, en el siguiente apartado profundizaremos algo más en esta cuestión.

En definitiva y como se puede comprobar, el proceso de diagnóstico es una tarea amplia y compleja en la que, en la mayoría de las ocasiones, es necesario contar con la intervención de distintos profesionales.

6. EL INFORME DE DIAGNÓSTICO

Como se ha visto, entre las diversas fases de un proceso de diagnóstico, hay una que implica la elaboración de un informe -oral y/o escrito- en el que se resumen, interpretan y, sobre todo, comunican las características esenciales y los resultados del proceso de exploración seguido. En este apartado pretendemos dar algunas orientaciones para la elaboración de informes diagnósticos, teniendo en cuenta que estos informes pueden adquirir varios formatos según las personas a las que van dirigidos, la finalidad con la que son elaborados y, por supuesto, el proceso de diagnóstico que hay detrás de ellos.

Las personas o grupos solicitantes o destinatarios de este informe pueden presentar características muy diferenciadas. Puede tratarse del propio sujeto sobre el que ha revertido el diagnóstico y éste, a su vez, puede ser una persona adulta, un niño o un adolescente. Puede tratarse también de la familia del sujeto del diagnóstico, de su jefe o superior, de un docente, de una institución, etc. Por lo tanto, una primera consideración que cabe hacer es que *el informe de un diagnóstico debe tener en cuenta la/s audiencia/s a la/s que va dirigido, adaptando su lenguaje y formato a las características de su/s destinatario/s, a fin de que pueda ser correctamente comprendido*.

Por otra parte, y como se ha planteado en los apartados precedentes, los procesos de diagnóstico educativo responden a una multiplicidad de posibles variables, situaciones, sujetos, etc. Así, el diagnóstico puede haberse centrado en un sujeto o en un grupo; puede, asimismo, haber sido aplicado en torno a unas pocas variables o a un conjunto amplio de ellas; puede haber respondido a diversas finalidades; etc. De esta forma, no será lo mismo un diagnóstico cuya finalidad es conocer el nivel de competencia curricular de un sujeto a fin de decidir las medidas educativas a tomar, que una evaluación de las actitudes hacia el maltrato en un grupo de madres y padres. El ámbito de aplicación del proceso de diagnóstico y sus propias características especiales delimitarán, por tanto, el enfoque y formato con que se realiza el informe.

No obstante, y de forma general, se pueden apreciar cuatro factores comunes en todo informe, independientemente de las características del proceso diagnóstico realizado. Estos factores son:

— Es una comunicación oral o escrita.

— Presenta una síntesis de la evaluación o diagnóstico efectuado.

— Indica adecuadamente las pautas y recomendaciones necesarias para la intervención educativa a desarrollar.

— Da respuesta a la demanda realizada, al motivo que suscita el proceso de diagnóstico/intervención.

En definitiva, el objetivo del informe es dar respuesta a los interrogantes que fueron planteados por el remitente del mismo. Puesto que cada informe puede adquirir matices peculiares, las pautas y recomendaciones que vamos a dar en este apartado deben ser adaptadas a cada contexto y situación, según las coordenadas en que se realiza el informe diagnóstico.

6.1. Cuestiones previas: formación, habilidad y ética en la elaboración del informe diagnóstico

Redactar un informe supone tener la habilidad suficiente no sólo para analizar, sintetizar e integrar todas las informaciones previamente recogidas, sino que también son necesarias las habilidades propias del lenguaje, que van a permitir que el destinatario pueda entender con precisión y claridad los mismos conceptos y terminología que en el informe se usan (JIMÉNEZ, 1997). Entre estas capacidades, podemos señalar las siguientes:

— Habilidad para encontrar el equilibrio entre los datos objetivos (resultados del proceso de exploración diagnóstica seguido) y la abstracción (o interpretación, integración y establecimiento de relaciones en la información recogida).

— Capacidad para suscitar y mantener el interés.

— Usar de forma inteligente las ilustraciones y ejemplos cuando resultan necesarios.

— Habilidad a la hora de mostrar la información relevante para la toma de decisiones.

Hay informes que se limitan a enunciar una serie de siglas (por ejemplo, abreviaturas de las pruebas utilizadas, abreviaturas de constructos y medidas psicológicas como CI, EM, etc.) y/o presentar una serie de cifras (como los resultados de las pruebas y tests aplicados). Ello imposibilita que sean comprendidos por otros profesionales o destinatarios (familia, educadores, empleadores, etc.) distintos del profesional que los emite.

No obstante estas consideraciones, los informes diagnósticos deben ser fieles a los datos recogidos y los procedimientos empleados. Así, el profesional que los confecciona debe tener siempre presente la línea que separa una correcta y adecuada interpretación de la información recogida, de juicios, valoraciones y afirmaciones de dudoso fundamento empírico. Como plantea JIMÉNEZ (1997:222), uno de los problemas fundamentales que se nos presenta a la hora de confeccionar un informe es «el poder llegar a formular, con la mayor precisión y claridad posibles, una descripción de la conducta de la persona que sea relevante e inteligible para tratar de resolver adecuada y fielmente los propósitos que motivaron el informe».

Además de las cualidades y habilidades a mostrar por el autor del informe, hay una serie de aspectos en los que es necesario hacer gala de nuestra profesionalidad en el

diagnóstico que hemos realizado y en el informe que se desprende del mismo. Todos estos aspectos son de gran importancia en tanto que entran dentro del terreno de la deontología profesional y dependen en gran medida de las prescripciones marcadas por los correspondientes códigos deontológicos (FERNÁNDEZ BALLESTEROS, 1992; SATTLER, 1982; etc.).

La confección de informes diagnósticos suele ser un punto clave en la mayoría de los códigos actualmente existentes. Puesto que en el próximo apartado profundizaremos en esta cuestión, nos limitamos por el momento a señalar algunas de las precauciones que, según estos códigos, se deben tener:

— Prudencia en el uso de términos que conllevan denotaciones peyorativas o que degeneran fácilmente en etiquetas (normal/anormal, inteligente/torpe, etc.).

— Garantía del derecho de confidencialidad, según el cual, los resultados de un diagnóstico o evaluación son de carácter íntimo y el profesional está obligado a no divulgar los mismos fuera de los límites concretos que la persona objeto del diagnóstico quiera establecer.

— Preservación del anonimato en las exposiciones públicas de los resultados de un proceso de diagnóstico.

— Claridad, rigurosidad e inteligibilidad de los informes, los cuales habrán de expresar los procesos desarrollados y las limitaciones de los mismos.

6.2. Propuesta de estructura del informe

Haremos referencia aquí a los apartados básicos a contemplar en un informe pedagógico, advirtiendo nuevamente que estos apartados dependerán en gran medida del tipo y características del diagnóstico realizado. Teniendo esto en cuenta, consideramos que los puntos más relevantes a contemplar en un informe de diagnóstico son los siguientes:

1. **Datos personales del sujeto.** Aquí se hace una breve presentación de los datos identificativos del sujeto (edad, profesión o situación académica, colegio...), así como de la fecha en que ha tenido lugar la exploración. A pesar de que existen otras opiniones al respecto (BUISÁN y MARÍN, 1987), en nuestra opinión, no es conveniente indicar el nombre completo de la persona/institución diagnosticada. Basta con usar las siglas o el nombre de pila o un código numérico que permita sólo al profesional del diagnóstico conocer qué informe pertenece a qué sujeto o institución. En todo caso, y especialmente si el informe lleva el nombre del sujeto evaluado, habrá que hacer constar de forma visible el carácter confidencial del mismo (SOBRADO y OCAMPO, 1997). En cambio, el informe sí debe identificar siempre al profesional que lo emite.
2. **Motivo del diagnóstico.** Descripción breve (no más de 7-10 líneas) de la situación que suscita el diagnóstico y de los campos que se han explorado (autoconcepto,

intereses, valores...). ¿Quién ha solicitado el diagnóstico?, ¿por qué?, ¿con qué finalidad se realiza éste?... La claridad de las ideas expresadas y la brevedad de la redacción son dos criterios muy importantes a tener en cuenta, ya que la finalidad de este apartado es que resuma de manera rápida la situación que ha motivado el proceso de diagnóstico y los factores que se han explorado.

3. **Proceso de exploración seguido.** Indicación de los procedimientos, estrategias e instrumentos de recogida de información que se han empleado. Se trata de presentar brevemente: el enfoque con que se ha realizado el diagnóstico, las estrategias empleadas, el por qué se han utilizado, para qué, en qué orden o secuencia, cuál ha sido el calendario de actuación, qué incidencias ha habido en la aplicación de cada procedimiento, etc. Todas las decisiones tomadas han de ser razonadas y justificadas. Se hará referencia a la información esencial sin florituras innecesarias. Los instrumentos que se han diseñado ad hoc para el diagnóstico en cuestión serán descritos brevemente pero no se incluyen aquí, sino en los correspondientes anexos. Se trata de que cualquier persona ajena al proceso de diagnóstico se haga una idea bastante aproximada de cómo se ha realizado éste, así como de las limitaciones o condicionantes que pueden existir en los resultados como consecuencia del proceso desarrollado.

4. **Presentación e interpretación de los resultados.** Diferenciando los distintos temas que han sido objeto de exploración, por ejemplo: descripción del contexto y trayectoria vital (académica y/o profesional) de ese sujeto, descripción de las puntuaciones y resultados en cada uno de los campos explorados (autoconcepto, estrategias de aprendizaje, intereses profesionales, etc.). En todo caso, se trata de exponer los resultados de forma detallada y sus implicaciones. Para ello, hay dos aspectos que conviene tener siempre presentes:

   — Es recomendable dividir la presentación de los resultados en diversos sub-apartados, de forma que se dedique uno de ellos a cada factor, rasgo, variable o conducta diagnosticada. Consideramos que, en la mayoría de los casos, esta opción es mejor que presentar los resultados en función de los instrumentos o estrategias aplicados. Esto último suele provocar que las informaciones estén dispersas a lo largo del informe, no permitiendo una lectura integrada de cada uno de los aspectos o dimensiones explorados. En cualquier caso, antes de presentar los resultados es necesario indicar de qué instrumento provienen (entrevista, test, observación...).

   — La interpretación de los resultados (derivados de la aplicación de estrategias tanto cuantitativas como cualitativas) ha de realizarse siempre correctamente. Mostrando exactitud, por un lado, en la interpretación de puntuaciones (por ejemplo, significado de un percentil, de una puntuación típica derivada...) y, por otro, en las implicaciones respecto a la dimensión evaluada (ateniéndonos al significado de la dimensión en cuestión según viene definida y operativizada por el instrumento utilizado). En la interpretación de los tests, habrá que hacer alusión a las

características de los baremos utilizados, especificando claramente si, en algún caso, se ha usado un baremo que no se corresponde adecuadamente con las características personales del sujeto.

5. **Síntesis valorativa y toma de decisiones.** Este apartado se enfoca a resumir y dar sentido a los resultados presentados. Se trata de plantear un resumen de las principales conclusiones, procurando establecer las oportunas relaciones entre los diversos hechos, situaciones o dimensiones exploradas. En él se pueden lanzar hipótesis explicativas de los resultados o plantear un modelo para entender y explicar la conducta observada. En todos los casos se presentará como «propuesta» del profesional del diagnóstico y se basará en datos reales suficientemente explicitados. Es muy importante formular los juicios valorativos que caracterizan a todo proceso de diagnóstico, explicitando cuáles son los referentes que se utilizan para emitir la valoración. Por ejemplo, si concluimos que un cierto alumno presenta necesidades especiales en relación a una determinada área del currículo, un referente importante será el conjunto de objetivos de esa área para el ciclo/curso del alumno en cuestión. Este apartado debe indicar claramente cuáles son las principales conclusiones que se pueden extraer y, sobre todo, cuáles son las principales necesidades de la persona en torno al tema objeto de exploración, explicitando asimismo una priorización de estas necesidades.
6. **Propuesta de intervención.** Aquí se trata de lanzar ideas sobre lo que sería conveniente desarrollar en una futura intervención. ¿En qué aspectos debe incidir ésta?, ¿qué sería bueno que esa persona/programa/institución contemplara o realizara?, ¿dónde podría buscar más información?, etc. Dependiendo del contexto en que se realiza el diagnóstico, este apartado podrá tomar diversas formas: desde un conjunto de recomendaciones generales, hasta una propuesta sistemática de intervención y su correspondiente evaluación y seguimiento (que, lógicamente, sería lo deseable en todos los casos).
7. **Bibliografía.** Si se ha mencionado o citado algún texto, la referencia de éste debe figurar aquí. Si están publicadas, las pruebas utilizadas deben figurar también en la bibliografía.
8. **Anexos.** En este apartado se presentarán, de forma ordenada y numerada, todas aquellas informaciones no imprescindibles para la correcta comprensión del informe pero que amplían los detalles que en éste solo se apuntan. Se deben incluir aquí las transcripciones de entrevistas, los registros observacionales, ejemplos de realizaciones del sujeto, los instrumentos diseñados específicamente por el profesional, etc.

6.3. Recomendaciones adicionales en la elaboración de informes

Además de los aspectos apuntados en los apartados precedentes, presentamos a continuación un listado de cuestiones a tener en cuenta en la confección del informe pedagógico. Este listado no pretende ser exhaustivo ya que, según las coordenadas y características de cada proceso concreto de diagnóstico, es posible que existan otros muchos aspectos a contemplar.

En líneas generales, en la elaboración de informes hay que prestar atención a:

— Explicitar siempre los datos originales sobre los que se basan las inferencias (MARTORELL, 1988).

— Remarcar con expresiones del tipo «tal vez», «probablemente», «parece que» nuestras afirmaciones cuando no estamos completamente seguros de ellas.

— Citar la fuente de información cuando algo no ha sido comprobado personalmente (por ejemplo, «según el padre...»).

— Es importante insistir en la presencia de aprendizajes, características y conductas más que en su ausencia, a no ser que se trate de las variables clave que suscitan el diagnóstico.

— Ajustar el lenguaje y la terminología al destinatario. El objetivo es que éste entienda perfectamente el contenido del informe. Cuando sean varios los destinatarios, si es necesario, habrá que modificar el lenguaje para cada uno de ellos. Por eso, a veces, la redacción de un informe conlleva la elaboración de dos o más versiones del mismo.

— En general, el lenguaje ha de ser claro, preciso y sin ambigüedades. Un informe lleno de ambigüedades y generalidades seguramente produzca una interpretación errónea. Asimismo, conviene evitar el uso de palabras vulgares y chabacanas que resten seriedad y rigurosidad al informe y diagnóstico realizados.

— Como regla general, no deben usarse abreviaturas (Pc, z, CI...).

— Usar una ortografía, gramática y puntuación correctas para evitar que se deteriore la imagen del profesional y del contenido del informe.

FERNÁNDEZ BALLESTEROS (1992) y BUISÁN y MARÍN (1987) sintetizan gran parte de estas recomendaciones en tres requisitos que han de reunir los informes diagnósticos:

— *Ser documentos científicos*, lo cual no implica que deban ser redactados con tecnicismos, sino que los apartados que contienen permitan identificar quién lo ha llevado a cabo y el procedimiento que ha seguido. Así, en un informe debe figurar: autor, datos de identificación del sujeto evaluado, motivo del diagnóstico, técnicas y procedimientos de recogida de información y de comprobación de resultados, el diagnóstico y las orientaciones.

— *Servir de vehículo para la comunicación,* que es lo mismo que decir que sea comprensible e inteligible para la persona que lo recibe. Se trata, por tanto, de

adecuar el lenguaje del emisor al del receptor. Ello no implica ocultar datos o hacerlo menos científico, sino emplear términos claros y sencillos que ejemplifiquen lo que queremos decir, evitando tecnicismos y expresiones ambiguas. Se evitarán también en lo posible las especulaciones y las interpretaciones «más allá de los datos». En la formulación de los diagnósticos se pondrá especial precaución en relativizar las predicciones, descripciones y explicaciones, bien dando los datos oportunos relativos al grado de probabilidad, o bien formulándolas con un lenguaje cuidadoso que evite las afirmaciones tajantes y el etiquetaje.

— *Ser útiles*. Esto dependerá de la concreción de las orientaciones. Así, informes muy completos y exhaustivos pueden quedar pobres si el lenguaje es ininteligible o la información no está bien organizada o no añade gran cosa a lo que ya se sabía. Un informe debe contener una buena guía práctica de lo que se debe hacer y cómo se debe hacer. Esto no implica que sea el profesional del diagnóstico quien siempre deba detallar el programa, pero sí le compete establecer las líneas directrices y las sugerencias necesarias.

Además de todas estas consideraciones relativas a la elaboración del informe, siempre que sea posible, es conveniente completar dicho informe con una entrevista. En los casos en que tenemos constancia de la dificultad de los destinatarios para entender el informe, esta estrategia resulta especialmente importante para asegurarnos la correcta y completa comprensión del mismo. Sin embargo, la entrevista es una oportunidad para el intercambio de ideas y opiniones en torno a los contenidos del informe. En ella, el destinatario puede expresar su punto de vista sobre las conclusiones de éste y pedir las aclaraciones que considere oportunas. En sí misma, esta entrevista es el primer paso para la intervención, ya que ayudará a que el sujeto tome conciencia de los resultados del diagnóstico y de la necesidad de poner en marcha un proceso de mejora.

6.4. Posibles criterios para valorar informes diagnósticos

Como resultado de las consideraciones realizadas sobre la confección de informes de diagnósticos educativos, proponemos aquí un conjunto de criterios que, de forma general, pueden servir para elaborar y/o valorar un informe diagnóstico:

— **Criterios formales.** Se refieren a:

- Estética y presentación adecuadas.
- Índice paginado.
- Inexistencia de errores ortográficos, gramaticales y tipográficos.
- Organización de anexos y correcta presentación de los mismos en el texto.

- Advertencia inicial (cuando se trata de informes realizados por estudiantes en prácticas).
- Brevedad en cuanto al número de folios: nunca más de los estrictamente necesarios.

— **Criterios de contenido.** Se refieren a:

- Claridad, brevedad y justificación del motivo que suscita el diagnóstico.
- Descripción de los instrumentos y la secuencia de aplicación, así como de circunstancias específicas que hayan podido incidir en los resultados.
- Adecuación de los procedimientos de análisis e interpretación.
- Presentación adecuada de los resultados.
- Integración y relación de los resultados (síntesis valorativa y priorización).
- Adecuación y coherencia de la propuesta de intervención.

— **Criterios instrumentales.** Como:

- Correcto diseño, selección y aplicación de los instrumentos y estrategias utilizados.
- Adecuación de las estrategias a la situación.
- Correcto uso e interpretación de los procedimientos de análisis (puntuaciones, categorías, etc.).

— **Criterios éticos/deontológicos.** Cuidar, entre otras cosas, aspectos como:

- La utilización del nombre completo del sujeto o de «etiquetas» inadecuadas para referirse al mismo.
- La identificación no autorizada de las personas, grupos o instituciones que han sido evaluadas.
- La tendencia a copiar de otros informes o de bibliografía.
- Las conclusiones precipitadas y no basadas en informaciones.
- Otras cuestiones que supongan el incumplimiento de los principios recogidos en los códigos deontológicos al uso.

Finalizamos este apartado dedicado al informe de diagnóstico añadiendo a todas las finalidades del informe ya comentadas una nueva función a la que éste también sirve. Así, como plantea GRANADOS (1997), el informe resulta de gran utilidad para las futuras acciones perfectivas del ejercicio profesional. Realizar el informe supone un proceso de distanciamiento y reflexión que permite que el profesional del diagnóstico se percate de los errores cometidos, de las deficiencias detectadas en el proceso, etc. Ello contribuirá a la mejora continua de su actuación profesional.

7. GARANTÍAS ÉTICAS Y DEONTOLOGÍA PROFESIONAL EN EL DIAGNÓSTICO Y LA EVALUACIÓN

Los profesionales de la pedagogía no disponemos, hoy por hoy, de un código deontológico aceptado y conocido por todos, que regule los aspectos clave de la conducta profesional. Se hace, por tanto, necesario reflexionar y explicitar cuáles son los valores que consideramos prioritarios en el ejercicio de nuestra actuación y sobre los que debe basarse nuestra intervención. Estos principios éticos nos permitirán formular una *concreción deontológica* o «conjunto sistemático que rige y dirige la conducta específica» (LÁZARO, 1998:449). La implicación ética del pedagogo es evidente, dada su estrecha relación con el ser humano dentro de un marco social y cultural determinado, por lo que el abordaje de la deontología profesional no puede limitarse a un conjunto de reflexiones puntuales.

Estas normas éticas se hacen aún más necesarias cuando nos referimos a las tareas de diagnóstico y evaluación, por las importantes repercusiones que sus métodos y resultados tienen. Algunos profesionales de la pedagogía, cuando ponen en práctica este tipo de tareas, piensan que su actuación es «neutral» por el hecho de usar un determinado tipo de instrumentos o proceso de evaluación que garantiza la objetividad de la información recogida. Como plantea BLASCO (1995), no podemos pensar que existe la neutralidad en la actuación profesional, puesto que, toda decisión técnica conlleva connotaciones éticas.

Otros colectivos profesionales con los que compartimos áreas de intervención tienen su correspondiente código deontológico que orienta y guía la conducta profesional. Tal es el caso de los psicólogos, cuyo código de conducta data de finales de los ochenta (COP, 1987). Asimismo, los profesionales de la orientación educativa y profesional también han elaborado un conjunto de normas éticas que deben regir la profesión del orientador (AIOSP, 1995; NCDA, 1987).

En el caso concreto de nuestro país, la deontología del quehacer pedagógico no ha recibido la atención que consideramos debe suscitar. A finales de la década de los ochenta, en el seno del III Encuentro Estatal de Pedagogos (Valdepeñas, Ciudad Real), se realizó un importante esfuerzo por sistematizar y definir las funciones propias de los pedagogos creándose su Estatuto Profesional (MORATINOS IGLESIAS, 1986). En ese mismo encuentro se llegó a formular un conjunto de normas éticas de nuestra profesión, gracias a la aprobación de un código deontológico profesional[2] (MORATINOS IGLESIAS, 1987 y 1992). Aunque este código ético supone la existencia de unas mínimas normas de conducta profesional, encontramos que aborda este tema desde una perspectiva muy general, no entrando en los detalles concretos de la actuación pedagógica, como sería necesario en un documento de esta índole. Otra problemática, no menos importante,

asociada a este código es su escasa difusión, lo que ha contribuido a que los profesionales de la pedagogía no sólo no lo tengan como referencia en su actuación profesional, sino que ni siquiera conozcan su existencia.

Además de lo anterior, consideramos que la sociedad española de finales de los años ochenta es muy diferente a la actual. No sólo hablamos de los cambios sociales en general, sino también de los importantes cambios y avances que se han producido en el terreno de la educación. Ello sin olvidar los cambios que afectan directamente a la profesión pedagógica. Así, han surgido nuevas especializaciones y campos de trabajo, nuevas metodologías y medios técnicos, nuevos recursos y nuevos perfiles profesionales que amenazan con contribuir a la ya considerable indefinición social del pedagogo.

Algunos autores han insistido en que las normas éticas no pueden formularse de manera permanente y universalista, sino teniendo presente una perspectiva multicultural (IBRAHIM, 1996; LÁZARO, 1998). La deontología educativa cambia con los tiempos, y también con los contextos culturales, al cambiar el sistema de valores, la ética y lo que se acepta por una sociedad como moralmente bueno.

En este marco, el código ético debe concebirse como una medida de autogestión responsable y profesional ante las demandas y exigencias sociales, y servir como defensa ante los problemas y necesidades educativas que actualmente se vienen planteando. En el terreno práctico, dicho código será muy útil, ya que marcará la pauta a seguir en los desacuerdos que pudieran surgir entre los mismos profesionales y en los conflictos con los poderes públicos y el público en general (BLÁZQUEZ, 1986). Por lo tanto, un código deontológico es un conjunto sistemático de normas y directrices que regula los deberes y obligaciones profesionales pero, a nuestro modo de ver, también es un marco general para suscitar la *reflexión colectiva* y la *conducta profesional reflexiva e informada*, sin pretender ser la panacea que resuelve felizmente todos los posibles conflictos profesionales.

7.1. Aspectos de la profesión pedagógica a incluir en un código deontológico

Ante la inexistencia de un código ético del pedagogo, adecuadamente consensuado y difundido entre el conjunto de los profesionales, en este apartado queremos reflexionar sobre los temas y aspectos que dicho código debería incluir, aportando nuestra propia propuesta al respecto. Para ello, resulta de gran utilidad hacer una revisión de los contenidos incluidos por los códigos deontológicos de profesiones afines (AIOSP, 1995; COP, 1987; NCDA, 1987) para, a partir de ahí, elaborar normas que recojan la idiosincrasia y peculiaridad del trabajo pedagógico en los múltiples campos en que éste se aplica. A este respecto, LÁZARO (1998) realiza un análisis de los códigos existentes en

el campo de la orientación que puede sernos de utilidad para conocer los contenidos que se suelen abordar en ellos, los cuales se refieren a:

1. *Aspectos generales de la relación pedagógica*. Tema que hace referencia a las condiciones que deben regir la relación entre el pedagogo y los destinatarios de su intervención.
2. *Confidencialidad de la información*. Este aspecto es de gran relevancia en los procesos diagnósticos y alude al uso restringido de la información resultante de cualquier diagnóstico o evaluación, así como a las precauciones a adoptar para impedir su conocimiento y difusión inadecuados.
3. *Actividades de responsabilidad profesional*. Se refieren a la actuación y dedicación del profesional de la pedagogía y a la asunción de los propios límites.
4. *Relación con otros profesionales y colegas*. Alude a los términos y coordenadas dentro de los cuales ha de darse la necesaria relación y cooperación con otros profesionales.
5. *Actividades de evaluación, valoración e interpretación de datos*. Este apartado, evidentemente, es de gran relevancia en los procesos de diagnósticos ya que a través de los diferentes artículos que se suelen incluir en él, se regulan todas las cuestiones relativas a la selección y aplicación de instrumentos, a la valoración e interpretación de la información recogida y a la elaboración de los correspondientes informes pedagógicos.
6. *Formación y supervisión profesional*. Apartado en el que se abordan todas las cuestiones relativas a la necesidad de actualización y formación permanente del profesional de la pedagogía como garantía de su competencia profesional.
7. *Uso de la información recogida en investigaciones y publicaciones*. Se refiere a la regulación de las condiciones en que los destinatarios participan en procesos de investigación pedagógica.
8. *Resolución de conflictos éticos*. Alude a la necesidad de conocer los códigos y normas existentes en la profesión y de ajustar a ellos el comportamiento profesional.

Una vez conocidos estos puntos, observamos que el código deontológico del pedagogo/a elaborado por la asamblea de Pedagogía en el III Encuentro Estatal de Pedagogos (ver anexo 1), resulta incompleto en algunos de los apartados anteriormente identificados. Así, dicho código se centra más en cuestiones relativas a los aspectos generales, actividades de responsabilidad profesional, relación con otros colegas y formación y supervisión profesional, tocando de forma tangencial aspectos clave como el relativo al secreto profesional, a las actividades de evaluación y al uso de la información recogida en investigaciones y publicaciones. Son estos, por otra parte, los aspectos más relevantes desde el punto de vista de esta obra, centrada en las técnicas e instrumentos de diagnóstico en educación.

A fin de ofrecer algunas pautas orientativas más concretas, en el anexo 2 presentamos una propuesta de normas éticas para la elaboración de un código deontológico del pedagogo. Dicha propuesta ha sido elaborada y consensuada gracias a la intervención de un conjunto bastante amplio de estudiantes y profesionales de la pedagogía en ejercicio. Es necesario advertir que se trata sólo y exclusivamente de una «propuesta» y no de un código deontológico aprobado y respaldado por ninguna asociación o colegio profesional. Por lo tanto, orientarse por estos preceptos y principios supone una opción personal y el incumplimiento de los mismos no conlleva consecuencias legales más allá de las que cada profesional o colectivo concreto quiera asumir.

Puesto que la propuesta se presenta íntegramente en el anexo 2 de este libro, aquí no nos referiremos a su contenido y nos limitaremos a destacar algunos aspectos que consideramos relevantes.

Uno de estos aspectos se refiere al uso de un vocabulario que no restrinja la amplitud de campos profesionales en los que el pedagogo puede trabajar. De esta forma, expresiones como «alumno/a», «profesores/as», «escuelas», etc. pueden llevar a pensar que el ámbito profesional del pedagogo se circunscribe al sistema educativo reglado. Puesto que la intervención pedagógica puede estar dirigida a otros colectivos e instituciones, se ha procurado utilizar un vocabulario que recoja esta diversidad («personas destinatarias de la intervención», «agentes educativos», «instituciones u organizaciones educativas», etc.).

Asimismo, otra cuestión que ha sido objeto de un intenso debate es la relativa al dilema que se produce entre la colaboración con otros profesionales y la necesidad de guardar el secreto profesional. Un ejemplo que puede ilustrar esta situación se encuentra en la figura del orientador de centros en la enseñanza secundaria. Tal y como está planteado su perfil, el orientador trabaja siempre en colaboración con otros profesionales. Si este profesional realiza un diagnóstico ante la demanda de uno o varios profesores, es evidente que debe compartir con ellos el resultado a fin de planificar y llevar a cabo las medidas educativas oportunas. Sin embargo, hasta cierto punto, esto contradice la obligación de no difundir los resultados de los procesos evaluativos. En la propuesta que hemos realizado (anexo 2), se ha procurado abordar esta cuestión en el artículo 12, especificando las condiciones en que se pueden comunicar los resultados de una evaluación en estos casos. Se puede observar que hemos optado por hacer un uso «consciente e informado» del secreto profesional, en tanto que se podrá compartir la información sólo cuando sea necesaria la «colaboración con fines de mejora» y, aún así, responsabilizando al pedagogo de informar adecuadamente a esas terceras personas de los términos en que deben hacer uso de dicha información.

Otra temática de importantes implicaciones es la referida al uso y custodia del material técnico estrictamente pedagógico. Este tema, que queda recogido en el actual artículo 10, también aborda una cuestión importante en el ejercicio de la pedagogía en determinados contextos en los que el trabajo pedagógico adopta un enfoque

interdisciplinar y colaborativo, que va en contra de la idea de que sea el pedagogo el que monopolice los instrumentos y materiales que sirven como medios para la evaluación y la intervención. Se plantea, por tanto, una duda importante que en el artículo citado se resuelve con la expresión «el pedagogo garantizará la custodia y uso legítimo de tales documentos», de forma que no se excluye la posibilidad de facilitar materiales de trabajo a otros profesionales (como, por ejemplo, al profesorado). Esto no implica que instrumentos como los tests se puedan utilizar por profesionales no cualificados, lo que entraría dentro de la misma idea de «uso legítimo» de los mismos.

A la vista de estas reflexiones, parece clara la idea de que la elaboración de un código deontológico es una tarea compleja que necesita del debate y consenso de toda la comunidad pedagógica. En todo caso y como decíamos anteriormente, un documento de este tipo no debe entenderse como *definitivo y universal*. Además de la evolución de la sociedad, la propia evolución y avance de la pedagogía lleva a que cualquier código ético tenga que ser necesariamente revisado, no sólo por la posibilidad de quedar obsoleto, sino también porque, en algunos casos, este desfase podría constituirse en una traba para el ejercicio de la pedagogía, en lugar de una herramienta necesaria para el avance de nuestra profesión.

Por otra parte, hemos de ser conscientes de que la sociedad cada vez es más sensible a los casos de negligencia e incompetencia profesional. Por citar algunos ejemplos de profesiones cercanas, en una revisión de la primera década de funcionamiento del código deontológico del psicólogo BATRES (2001) plantea que la mayor parte de las denuncias e infracciones que se han producido tienen que ver con la *competencia profesional*. Ésta se refiere, entre otras cosas, a la formación y especialización necesarias para la utilización de los métodos, instrumentos, técnicas y procedimientos que adopte en su trabajo. Se detectan así muchos casos de inadecuada selección y utilización de las técnicas, falta de método y planificación en la intervención e informes mal estructurados y presentados. Asimismo, se está incrementando el número de denuncias en la realización de determinados informes solicitados por instancias administrativas y judiciales (BATRES, 2001; DEL RÍO, 2000).

Observando el contenido de estas infracciones, por otra parte, podemos inferir que las tareas relacionadas con el *diagnóstico* en general y con el *uso de técnicas e instrumentos para el mismo* en particular, constituyen una de las cuestiones clave y de mayor trascendencia y relevancia a abordar en un código deontológico del pedagogo. Como conclusión, para promover una práctica de calidad en este campo, resulta necesaria una sólida formación del profesional de la pedagogía en estos temas, es decir, en el diagnóstico educativo y sus técnicas y procedimientos.

8. CONSIDERACIONES SOBRE LA SELECCIÓN DE INSTRUMENTOS PARA EL DIAGNÓSTICO Y LA EVALUACIÓN

En los siguientes capítulos de este libro abordaremos el análisis de las principales técnicas e instrumentos para el diagnóstico en educación, si bien en estas páginas introductorias queremos reflexionar sobre estas estrategias de una forma más global.

Si tuviéramos que definir lo que consideramos como un buen profesional del diagnóstico educativo en lo que respecta al uso de las diversas técnicas, más que aquel profesional que conoce y usa muchas de ellas, o aquel otro que domina con alto grado de excelencia un reducido conjunto de las mismas, nos decantaríamos por el que sabe usar la estrategia más adecuada para cada caso o situación y, además, en la medida de lo posible, procura basar sus diagnósticos en el uso combinado de varias estrategias.

Pese a que, históricamente, ciertas técnicas han gozado bien de la aprobación general, bien del más absoluto rechazo, cuando hablamos de «técnicas e instrumentos de diagnóstico» no hacemos sino referirnos a las «herramientas» de trabajo del profesional de la pedagogía. Por ello, no cabe calificar las técnicas como «buenas» o «malas», sino distinguir entre procesos de diagnósticos correctamente planificados y realizados, y procesos de diagnósticos que se alejan de una práctica profesional de calidad. Haciendo un símil con el oficio de carpintero, un martillo o una sierra pueden usarse eficientemente para construir un excelente armario, o también de forma torpe y chapucera para convertir una madera de calidad en una parodia de ropero. E incluso, cabe utilizar tanto el martillo como la sierra para hacer daño a un ser humano. Por lo tanto, en nuestra opinión, no hay instrumentos mejores o peores, sino profesionales que saben o no usarlos convenientemente. En esta línea se expresa MUÑIZ (1997:317):

> «En esto ocurre con los tests lo mismo que con cualquier otra tecnología, sea del campo que sea, su potencialidad para causar daño proviene más de su uso por gentes sin preparación o principios que de sus propias técnicas».

Un ejemplo de esta confusión entre la técnica y el proceso de diagnóstico es la que se ha dado con el uso de los tests. Como se ha indicado en apartados anteriores, determinados enfoques de evaluación partieron de la identificación del diagnóstico con la testación, es decir, consideraban que todo proceso de diagnóstico consistía en la aplicación de tests, y que tan sólo estos instrumentos garantizaban la objetividad y cientificidad de tal proceso. Sin embargo, los tests no son instrumentos aplicables en todos los casos, además de presentar importantes sesgos y no garantizar siempre de forma adecuada la fiabilidad y validez de la medición que aportan. La validez predictiva de estos instrumentos es una de las cuestiones más problemáticas y en la que se han cometido los mayores excesos. En el campo educativo, por ejemplo, se ha demostrado la ineficacia de estos instrumentos para predecir el futuro éxito escolar. Así, una buena o

mala ejecución en las tareas propuestas en una determinada prueba no implica necesariamente una buena o mala ejecución en las tareas escolares y, mucho menos, el éxito o fracaso académico, que depende de otros muchos factores (hábitos de estudio, adaptación al contexto escolar, influencia de la familia, metodología didáctica, materiales de enseñanza, etc.).

Como consecuencia de estas prácticas diagnósticas, se produjo un fuerte rechazo al uso de los tests. Este rechazo no sólo afectó a la sociedad en general (padres y maestros descontentos ante sus resultados), sino que también afecta hoy día a muchos profesionales de la pedagogía, que rechazan de plano las pruebas estandarizadas y critican su uso con vehemencia, considerando como inadecuada cualquier medición normalizada del comportamiento.

En nuestra opinión, si bien los tests no son la panacea en materia de evaluación, consideramos que son de utilidad para la medición de determinadas características y nos aportan una ayuda válida allí donde otros instrumentos no llegan. Sin embargo, estamos de acuerdo con que se han usado de forma poco correcta, llevando sus resultados más allá de lo que está permitido por las normas estadísticas al uso y por el propio sentido común. Por ello, consideramos que tan fuera de lugar está la fe ciega en las virtudes de los tests como una reacción desmedida en contra de ellos.

La cuestión no está, por tanto, en no usar las pruebas estandarizadas, sino en reconocer las limitaciones que tienen y obrar en consecuencia. Como profesionales de la pedagogía, hemos de reconocer que el uso de tests en los procesos diagnósticos nos confiere una «autoridad» a la que resulta difícil renunciar. Nos permiten cuantificar dimensiones, comparar resultados de varios sujetos y, por supuesto, es lo más parecido a trabajar con «aparatos» de diagnóstico que podemos encontrar en nuestra profesión. Sólo su «apariencia» física ya es impresionante: el cuadernillo con las preguntas e instrucciones (que, en ocasiones, lleva códigos ininteligibles para el sujeto), la hoja de respuestas (muchas veces ya preparada para la corrección mecánica), puntuaciones numéricas, normas y baremos, y toda una parafernalia que conduce a pensar que ofrecen garantías de mayor cientificidad (objetividad, precisión de la medida, interpretaciones categóricas a las que conducen, etc.). Esta «apariencia» nos lleva a creer en la falacia de que son instrumentos investidos de un poder especial que sólo el profesional puede usar e interpretar. Una de las consecuencias de esto es que la persona destinataria del diagnóstico no tiene posibilidades de intervención, puesto que se mueve en un terreno en el que el experto es el pedagogo.

En cualquier caso, los tests no son los únicos instrumentos que tienen limitaciones, si bien son los que más críticas han recibido. En ocasiones, hemos comprobado que muchos profesionales rechazan los tests y cualquier técnica estructurada de recogida de información por un temor irracional a «trabajar con números». Acuden así al uso masivo de las técnicas de carácter más cualitativo. Esto, a nuestro juicio, encierra dos nuevos peligros de los que no siempre somos conscientes:

— Por un lado, poder cuantificar las variables objeto de nuestro interés es necesario en ocasiones para tener una idea de la magnitud de un fenómeno (el grado de incidencia del mismo en una persona o el porcentaje de personas a las que afecta, por poner algunos ejemplos). No podremos hablar de «diagnósticos completos» si rehuimos constantemente de todo procedimiento que implique cuantificación.

— Por otro, si bien el uso de ciertas técnicas requiere unos conocimientos y competencias previas en estadística, no es menos cierto que las técnicas cualitativas también necesitan una sólida formación previa. Si este requisito no se da, cometemos el mismo error aplicando un test sin saber interpretarlo correctamente, que utilizando una historia de vida o una entrevista abierta que no sabemos dinamizar, aplicar e interpretar.

Por lo tanto, pese a las aportaciones que las técnicas cualitativas pueden realizar a los procesos diagnósticos, hemos de tener en cuenta que, al igual que los tests y otros procedimientos, tienen sus limitaciones y son susceptibles de ser utilizadas indiscriminada e incorrectamente. Un profesional del diagnóstico que quiera utilizar este tipo de estrategia no debe acudir a ellas por su falta de formación en el manejo de otro tipo de pruebas, ya que las técnicas cualitativas exigen no sólo el conocimiento de los procedimientos a utilizar, sino también la experiencia previa y el «rodaje» en el uso de los mismos, si se quiere sacar el máximo partido de ellas. Como plantea GOLDMAN (1992), las diferentes competencias necesarias para la aplicación de pruebas estandarizadas y de técnicas cualitativas no es una cuestión de cantidad, sino de tipos de habilidades diferentes para su puesta en práctica.

En resumen, el conocido refrán que dice que «en la variedad está el gusto» es quizá una buena forma de expresar la conveniencia de usar diferentes técnicas en todo proceso de diagnóstico y evaluación. Ahora bien, dicho proceso no debe convertirse en un *tótum revolútum* en el que no existe una planificación adecuada de medios y fines. Cada técnica sirve a unos propósitos. Conocer para qué son útiles, ser conscientes de sus limitaciones y saber adaptarlas a cada situación concreta son las competencias que deben mostrar los buenos profesionales de la pedagogía.

Referencias bibliográficas

AIOSP (1995). Normas éticas de la AEOP. *Revista de Orientación Educativa y Vocacional*, *6* (10), 150-155.

ÁLVAREZ, M. (1991). Modelos y programas de intervención en orientación. *Actas del VIII Encuentro de la Asociación Coordinadora de Pedagogía*. Albacete: Abril.

ÁLVAREZ, M. (1995). *Orientación profesional*. Barcelona: Cedecs.

ÁLVAREZ ROJO, V. (1984). *Diagnóstico pedagógico*. Sevilla: Alfar.

ANAYA, D. (1994). *Introducción al diagnóstico en orientación. Bases conceptuales y metodológicas*. Madrid: Sanz y Torres.

BATANAZ, L. (1996). *Investigación y diagnóstico en educación: una perspectiva psicopedagógica*. Málaga: Aljibe.

BATRES, C. (2001). Deontología profesional: fin de una etapa. *Papeles del Psicólogo, 78*, 29-31.

BLASCO, T. (1995). *La ética en la investigación y en la práctica psicológica*. Barcelona: Ariel.

BLÁZQUEZ, N. (1986). Deontología de la educación. *Revista Española de Pedagogía*, *174*, 483-500.

BUISÁN, C. y MARÍN, M. A. (1984). *Bases teóricas y prácticas para el diagnóstico pedagógico*. Barcelona: PPU.

BUISÁN, C. y MARÍN, M. A. (1986). *Tendencias actuales en el diagnóstico pedagógico y la orientación*. Barcelona: Laertes.

BUISÁN, C. y MARÍN, M. A. (1987). *Cómo realizar un diagnóstico pedagógico*. Barcelona: Oikos-tau.

BRUECKNER, L. J. y BOND, G. L. (1971). *Diagnóstico y tratamiento de las dificultades de aprendizaje*. Madrid: Rialp.

COLEGIO OFICIAL DE PSICÓLOGOS (1987). *Código deontológico del psicólogo.* Colegio Oficial de Psicólogos. Madrid.

CASTILLO, S. (1994). Diagnóstico-evaluación: cooperación y complementariedad. *Revista de Investigación Educativa*, *23*, 640-642.

DEL RÍO, C. (2000). Informes de parte en conflictos matrimoniales: implicaciones deontológicas. *Infocop, 10*, 15-20.

FERNÁNDEZ BALLESTEROS, R. (1986). *Psicodiagnóstico.* Madrid: UNED.

FERNÁNDEZ BALLESTEROS, R. (1992). La comunicación de los resultados del proceso: el informe. En R. Fernández Ballesteros (dir.), *Introducción a la evaluación psicológica,* Vol. 1 (pp. 87-108). Madrid: Pirámide.

GARANTO, J. (1997). Investigación y prospectiva en educación especial: el diagnóstico. *Educar, 21*, 85-103.

GARCÍA JIMÉNEZ, E. (1994). ¿Diagnóstico vs. evaluación? *Revista de Investigación Educativa*, *23*, 584-587.

GARCÍA NIETO, N. (1990). El diagnóstico pedagógico y la orientación educativa unidos en un mismo proceso. *Bordón, 42* (1), 73-78.

GARCÍA VIDAL, J. y GONZÁLEZ MANJÓN, D. (1992). *Evaluación e informe psicopedagógico*. Madrid: EOS.

GIL FERNÁNDEZ, P. (1983). Diagnóstico. En Santillana (ed.). *Léxicos de Ciencias de la Educación*. Madrid: Santillana.

GOLDMAN, L. (1992). Qualitative assessment: an approach for counselors. *Journal of Counseling and Development, 70*, 616-621.

GRANADOS, P. (1997). *Diagnóstico pedagógico. Addenda*. Madrid: UNED.

IBARRA, M. S. (1999). *Diagnóstico en educación.* Proyecto Docente. Cádiz: Universidad de Cádiz.

IBRAHIM, F. (1996). A multicultural perspective on principles and virtues ethics. *The Counseling Psychologist*, *24* (1), 78-85.

JIMÉNEZ, F. (1997). El informe psicológico. En G. Buela-Casal y J.C. Sierra (Dirs.), *Manual de evaluación psicológica* (pp. 221-238). Madrid: Siglo XXI.

LÁZARO, A. (1986). *Orientación y educación especial*. Madrid: Anaya.

LÁZARO, A. (1990). Problemas y polémicas en torno al diagnóstico pedagógico. *Bordón, 42* (1), 7-15.

LÁZARO, A. (1997). Variables relevantes para el diagnóstico en situaciones ordinarias en los diferentes niveles del sistema educativo. En H. Salmerón (coord.), *Diagnosticar en educación* (pp. 163-182). Granada: FETE-UGT.

LÁZARO, A. (1998). Ética profesional y códigos deontológicos. En R. Bisquerra (coord.), *Modelos de orientación e intervención psicopedagógica* (pp. 449-456). Barcelona: Praxis.

MAGANTO, C. (1995). *Psicodiagnóstico infantil. Aspectos conceptuales y metodológicos*. Bilbao: Universidad del País Vasco.

MAGANTO MATEO, J. (1996). *Diagnóstico en educación*. Bilbao: Servicio de Publicaciones de la Universidad del País Vasco.

MARÍN, M. A. (1997). Conceptualización del diagnóstico para una orientación educativa de calidad. En H. Salmerón (coord.), *Diagnosticar en educación* (pp. 17-46). Granada: FETE-UGT.

MARTÍ QUIRÓS, J. L. (1982). *Las deficiencias en el aprendizaje. Tipos y causas*. Barcelona: Nau Llibres.

MARTÍNEZ GONZÁLEZ, R. A. (1993). *Diagnóstico pedagógico*. Oviedo: Servicio de Publicaciones de la Universidad de Oviedo.

MARTÍNEZ GONZÁLEZ, R. A. (1994). ¿Diagnóstico pedagógico o evaluación educativa? Hacia una clarificación conceptual y terminológica. *Revista de Investigación Educativa, 23*, 623-630.

MARTORELL, M. C. (1988). *Técnicas de evaluación psicológica (vol. 1)*. Valencia: Promolibro.

MINISTERIO DE EDUCACIÓN Y CIENCIA (1990). *Orientación educativa e intervención psicopedagógica*. Serie Documentos. Madrid: MEC.

MORATINOS IGLESIAS, J. (1986). Las funciones del pedagogo y su estatuto profesional. *Revista de Ciencias de la Educación*, *128*, 515-522.

MORATINOS IGLESIAS, J. (1987). El código deontológico profesional del pedagogo. *Revista de Ciencias de la Educación*, *130*, 239-244.

MORATINOS IGLESIAS, J. (1992). La profesión de pedagogo al servicio de la calidad educativa. *Revista de la Asociación Coordinadora Estatal de Pedagogía*, *18*, 3-10.

MUÑÍZ, J. (1997). Aspectos éticos y deontológicos de la evaluación psicológica. En A. Cordero (coord.), *La evaluación psicológica en el año 2000* (pp. 307-345). Madrid: TEA.

NATIONAL CAREER DEVELOPMENT ASSOCIATION (1987). *National Career Development Association Ethical Standars*. New Orleans, LA: NCDA Board of Directions. (También publicado en Spokane, A.R. (1991). *Career intervention* (pp. 312-319). Englewoods Cliffs, NJ: Prentice-Hall).

ORDEN HOZ, A. DE LA (1969). La evaluación del rendimiento educativo y la calidad de la educación. *Revista de Educación*, 206.

PARRA, J. (1996). *Diagnóstico en Educación*. Murcia: Instituto de Ciencias de la Educación.

PÉREZ JUSTE, R. (1989). Fundamentos. En R. Pérez Juste y J.L. García Ramos, *Diagnóstico, evaluación y toma de decisiones* (pp. 17-70). Madrid: Rialp.

PÉREZ JUSTE, R. (1990). Recogida de información en el diagnóstico pedagógico. *Bordón, 42* (1), 17-29.

PÉREZ JUSTE, R. (1994). ¿Por qué diferenciar los términos evaluación y diagnóstico? *Revista de Investigación Educativa*, *23*, 588-591.

RODRÍGUEZ ESPINAR, S. (1982). Diagnóstico y predicción en orientación. *Revista de educación, 270*, 113-140.

RODRÍGUEZ ESPINAR, S. (coord.), ÁLVAREZ, M., ECHEVERRÍA, B. y MARÍN, M.A. (1993). *Teoría y práctica de la orientación educativa.* Barcelona: PPU.

SANZ ORO, R. (1990). El diagnóstico pedagógico desde la perspectiva de programas de intervención en la orientación educativa. *Bordón, 42* (1), 65-72.

SATTLER, J. M. (1982). *Evaluación de la inteligencia infantil y habilidades especiales*. México: El Manual Moderno.

SILVA, F. (1982). *Introducción al psicodiagnóstico*. Valencia: Promolibro.

SILVA, F. (1989). *Evaluación conductual y criterios psicométricos*. Madrid: Pirámide.

SOBRADO, L. y OCAMPO, C. (1997). *Evaluación psicopedagógica y orientación educativa*. Barcelona: Estel.

CAPÍTULO

2

INTRODUCCIÓN AL DISEÑO Y APLICACIÓN DE PROCEDIMIENTOS OBSERVACIONALES

1. LA OBSERVACIÓN COMO ESTRATEGIA DIAGNÓSTICA

1.1. Concepto y características

Se puede decir que la observación es el método por excelencia para recabar información en un proceso de diagnóstico. Así lo sugieren autores como SIERRA BRAVO (1983), quien afirma que la observación, entendida en sentido amplio, engloba todos los procedimientos utilizados en las ciencias sociales. De hecho, cuando aplicamos un cuestionario, una prueba objetiva o cualquier otro procedimiento, lo que hacemos no es sino someter a los sujetos a un estímulo y *observar* cuál es su respuesta al mismo para, a partir de esta respuesta, inferir alguna característica o rasgo en la persona objeto de nuestro diagnóstico.

La observación es, por tanto, una estrategia muy adecuada y útil, no sólo para los procesos evaluativos y diagnósticos, sino también para la investigación educativa en general. Ahora bien, existen múltiples situaciones cotidianas en las que estamos observando. Cuando pasamos por un escaparate y comparamos los productos y sus precios, cuando entramos en una sala y buscamos a alguien, cuando vamos a cruzar la calle y comprobamos que no pasan coches... En todas estas situaciones estamos haciendo uso de la acepción más común y genérica del término *observación* que, según el diccionario es, *examinar profundamente*. Ahora bien, esta observación espontánea, tácita, casual y, en muchas ocasiones, superficial, no responde al concepto de observación que vamos a utilizar en este capítulo, ya que aquí nos referiremos constantemente a una forma de observación que se caracteriza por ser científica.

¿Qué es lo que hace que una observación sea *científica*? Responder a esta pregunta no es nada fácil porque nos remite al propio concepto de *ciencia,* cuyo significado es diferente según la posición que cada comunidad científica defiende. Así, dado el carácter social del concepto de *ciencia*, consideramos que fijar qué es ciencia es una tarea que sólo puede realizarse de forma relativista, como plantea DENDALUCE (1988), es decir, teniendo en cuenta los valores de la comunidad científica.

No obstante lo anterior, ANGUERA (1982) hace una aproximación bastante amplia y genérica a las características que confieren a la observación el carácter de técnica científica. Estas características son:

— Servir a un objetivo ya formulado de investigación (o diagnóstico).

— Ser planificada sistemáticamente.

— Ser controlada y relacionada con proposiciones más generales, en vez de ser presentada como una serie de curiosidades interesantes.

Hay un cuarto rasgo, señalado por ANGUERA (1982), que creemos resulta interesante analizar con mayor detalle. Así, «estar sujeta a comprobaciones de validez y fiabilidad» sería también, en opinión de esta autora, un requisito importante para considerar científica la observación. No obstante, cualquier comprobación de la validez y la fiabilidad presupone ya un tipo de observación que es susceptible de ser cuantificada y, entre las diversas modalidades de observación que hoy día usamos en el campo educativo, encontramos algunas que no se utilizan con esta finalidad.

Por tanto, conviene matizar que, aunque estamos de acuerdo con el requisito planteado por ANGUERA (1982), nos parece necesario adoptar una perspectiva más amplia en su formulación. De este modo, deberíamos decir que la observación científica «debe estar sujeta a comprobaciones de calidad en sus características técnicas», siendo diferentes los criterios y procedimientos que utilizamos para comprobar esa calidad en función del tipo de estrategia observacional. En el apartado final de este capítulo, el lector podrá encontrar algunas reflexiones sobre los criterios que se pueden aplicar para estimar la calidad de una técnica o instrumento diagnóstico.

COLÁS (1994) también se refiere a las características que hacen de la observación una actividad científica, haciendo mención de alguna de las ideas aportadas por ANGUERA (1982). Así, en opinión de esta autora, la observación ha de ser:

— Intencional, lo que implica partir de una delimitación previa de lo que se va a observar, a quién se observará, en qué condiciones, etc.

— Estructurada, en tanto que vinculada a teorías e hipótesis.

— Controlada, es decir, objetiva, comprobable.

Queda, por tanto, clara la idea de que aquí hablamos de la observación como un proceso *deliberado* y *sistemático* que ha de estar orientado por una pregunta, propósito o problema. Este problema o propósito es el que da sentido a la observación en sí, y el que determina aspectos tales como qué se observa, quién es observado, cómo se observa, cuándo se observa, dónde se observa, cuándo se registran las observaciones, qué observaciones se registran, cómo se analizan los datos procedentes de la observación o qué uso se da a los mismos (EVERTSON y GREEN, 1989). Además, siguiendo a estas mismas autoras, el propósito de la observación está relacionado con la teoría, las creencias, los presupuestos y/o experiencias previas de la persona que efectúa la observación. Estos factores conforman el *marco de referencia* del observador e inciden en las decisiones que se toman, así como en el proceso observacional.

«El observador es el primer instrumento de observación», afirman EVERTSON y GREEN (1989:309), ya que su sistema perceptual es el primer instrumento que éste utiliza. Dicho de otra forma, el objeto seleccionado, el marco de referencia del observador y el propósito de la observación, entre otros factores, ejercen su influencia sobre lo que se percibe, registra y analiza. Independientemente de ello, el observador puede servirse de

determinadas herramientas o instrumentos que le ayudarán a focalizar o guiar su observación (un sistema categorial, notas de campos, etc.).

Conviene ahora que nos adentremos más profundamente en el concepto de observación y planteemos qué es lo que diferencia a los procedimientos observacionales de otros procedimientos que usamos habitualmente en el diagnóstico y la evaluación. Así, según ANGUERA (1982:20), observar es «advertir los hechos como espontáneamente se presentan y consignarlos por escrito».

Otra definición de la misma autora (ANGUERA, 1988:7) es la que concibe la observación como un procedimiento encaminado a articular una percepción deliberada de la realidad manifiesta con su adecuada interpretación, captando su significado de forma que, mediante un registro objetivo, sistemático y específico de la conducta generada de forma espontánea en un determinado contexto, y una vez se ha sometido a una adecuada codificación y análisis, nos proporcione resultados válidos dentro de un marco específico de conocimiento.

También SIERRA BRAVO (1983:213) nos ofrece una definición de observación como:

> «la inspección y estudio realizado por el investigador, mediante el empleo de sus propios sentidos, especialmente de la vista, con y sin ayuda de aparatos técnicos, de las cosas y hechos de interés social, tal como son o tienen lugar espontáneamente en el tiempo en que acaecen y con arreglo a las exigencias de la investigación científica».

Entre las notas o características que este autor atribuye a la observación, destacamos las siguientes:

— Ser un procedimiento de recogida de datos que se basa en lo percibido por los propios sentidos de la persona que realiza la investigación o el diagnóstico.

— Consistir en el estudio de fenómenos que existen natural y espontáneamente.

— Ser un examen de fenómenos o acontecimientos actuales tal como son o tienen lugar en la realidad del momento presente.

ANGUERA (1995a:527) también hace referencia a la observación directa, a la «que podemos considerar como la observación en sentido estricto». La observación directa requiere conductas perceptibles mediante nuestros órganos sensoriales. A un nivel inicial podemos, por tanto, decir que *observar* es en primer lugar *percibir*. Ahora bien, «si en el proceso observacional sólo actuara la percepción (visual, o auditiva, u otras), se correría el riesgo de pretender que un medio técnico de registro pudiera sustituir al observador humano, con lo cual el registro estaría compuesto de señales vacías de contenido. Es preciso que los datos resultantes del mecanismo representacional (informaciones percibidas) sean interpretados adecuadamente, de forma que se confiera un determinado sentido a lo percibido» (ANGUERA, 1995 a:530).

Bajo estas premisas, Mucchielli (1974:11) ha formulado la *ecuación fundamental de la observación,* que vendría a expresarse del siguiente modo:

$$O = P + I + C_P - S$$

Donde *O* es observación, *P* percepción, *I* interpretación, C_p conocimiento previo y *S* sesgos. La ecuación viene a denotar que la observación implica un adecuado equilibrio entre la percepción, la interpretación y el conocimiento previo, aunque siempre existan determinados sesgos que le pueden afectar.

La observación, por tanto, como estrategia de diagnóstico, se caracteriza por el hecho de recoger información desde la perspectiva perceptual e interpretativa de la persona que observa. Además de los procesos perceptuales e interpretativos, en toda observación está implicado algún tipo de registro o consignación por escrito de los hechos, conductas y fenómenos que se perciben durante la misma.

En cuanto a sus aplicaciones potenciales, muy pocas variables son incompatibles con la observación como técnica de recogida de información, especialmente en el campo educativo. PÉREZ JUSTE (1990) señala aquéllas que, debido a su complejidad, se prestan a ello en mayor medida que otras: el ámbito de la relación humana, el análisis de los procesos y la influencia de los contextos en la conducta.

Asimismo, y siguiendo a este autor, la observación como técnica puede utilizarse como complemento y contraste de la información recibida por otras técnicas (pruebas, escalas, cuestionarios...) pero también como técnica fundamental o primaria. Cuando actúa como complemento, la observación va a resultar muy útil a la hora de comprobar con conductas reales la veracidad de las afirmaciones realizadas en pruebas escritas. Al mismo tiempo, permite contrarrestar la situación de artificialidad presente en este tipo de pruebas.

Las ventajas e inconvenientes de la observación como estrategia diagnóstica han sido abordadas por ANGUERA (1982). Esta autora propone un conjunto de ventajas de esta técnica, entre las que resaltamos las siguientes:

— Hace posible obtener la información tal como ocurre, es decir, las técnicas de observación aportan datos que atañen directamente a situaciones de comportamiento típicas.

— Muchas de las conductas y situaciones a observar pasan desapercibidas para los actores, mientras que son percibidas por los observadores.

— Permite acceder a sujetos que no pueden proporcionar informes verbales (por ejemplo, las criaturas pequeñas que no saben hablar).

— Es adecuada para cuando los sujetos o grupos de sujetos del diagnóstico manifiestan resistencia al mismo. Así, dado que la observación solicita en menor medida la cooperación activa de los sujetos, puede encontrar menos resistencia que otros procedimientos de recogida de información que requieren mayor participación.

Entre sus limitaciones, la misma autora (ANGUERA, 1988) señala la imposibilidad de predecir la ocurrencia espontánea de un suceso que hace que no siempre sea posible estar presente para proceder a la observación. Además, incluso sucesos habituales pueden resultar difíciles de observar en tanto que interfieren factores ocultos. Por otra parte, la posibilidad práctica de hacer observación se halla limitada por la duración de los sucesos o el contexto en el que tienen lugar (evidentemente, no podemos estar presentes en muchas situaciones privadas o íntimas de la vida humana). Por último, se achaca a la observación el hecho de que sus datos no se pueden cuantificar, lo cual, en opinión de esta autora, es erróneo porque en muchos casos sí que es posible esta cuantificación.

1.2. Modalidades de observación

Conviene empezar este apartado estableciendo una diferencia inicial entre la *observación* y su *registro*. Así, bajo nuestra perspectiva, la observación hace referencia a la percepción e interpretación de la realidad por parte de la persona que la realiza. Ahora bien, esa percepción de la realidad está incompleta si no se produce algún tipo de registro en el que quede constancia de lo observado.

BLANCO y ANGUERA (1991) plantean que todo proceso de observación supone una actividad de codificación o de transformación de la información bruta seleccionada gracias a un código para poder ser transmitida a otros. En un sentido más restringido, la observación designará el resultado codificado del propio acto de observar. En un sentido más amplio, la observación será el resultado codificado del acto de observar seguido del acto de interpretar, lo cual supondrá la referencia a un marco teórico o explicativo del fenómeno objeto de observación.

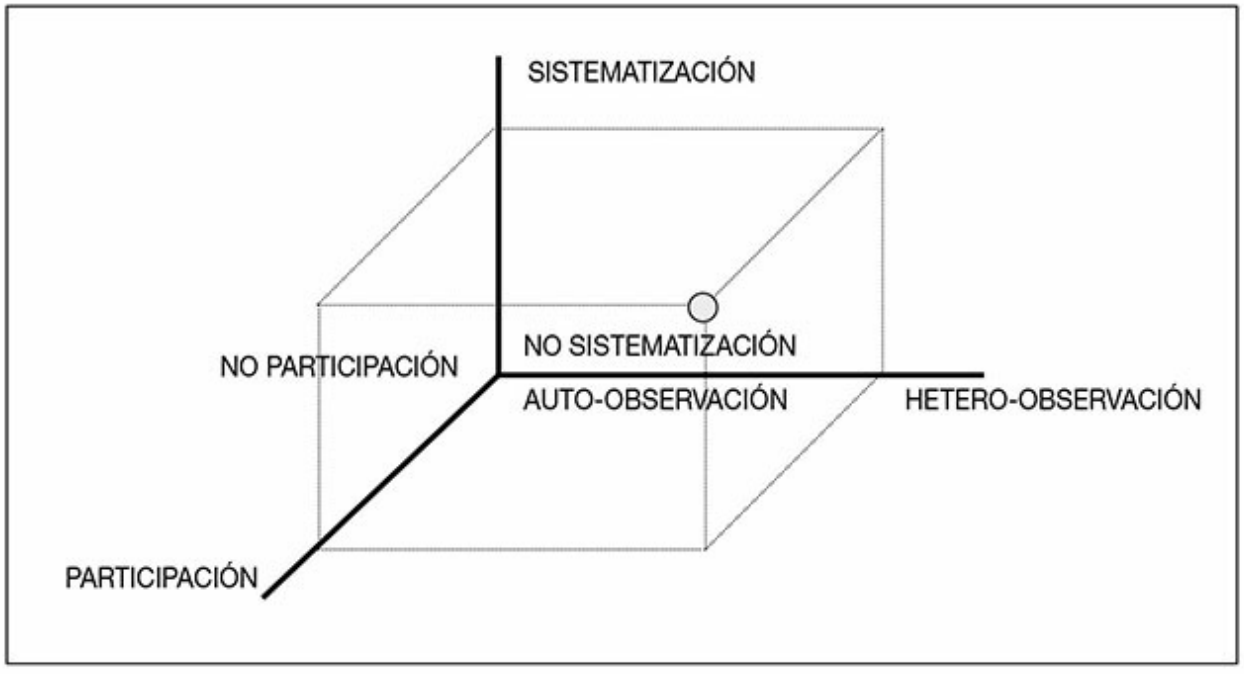
Hecha esta diferenciación, cabe distinguir también entre *modalidades de observación* y *tipos de registros.* Las primeras hacen alusión a la forma de enfocar la observación, fruto del marco teórico de referencia que se utiliza. Se puede, así, enfocar la observación de forma que el observador participe en los fenómenos que observa (observación participante), o de forma que no se dé esta participación (observación no participante); se puede optar por delimitar de forma muy precisa lo que vamos a observar y cómo vamos a hacerlo antes de proceder a realizar la observación (observación cerrada o sistematizada), o se puede realizar la observación partiendo sólo de ideas muy generales (observación abierta o no sistematizada); se puede observar el comportamiento de otras personas (hetero-observación), o ser el propio observador el objeto de la observación (auto-observación). Todas estas opciones engloban diferentes modalidades de observación.

Sin embargo, además de decidir el marco general y enfoque desde el que se realiza la observación, es preciso decidir cómo se va a registrar la información recabada a través de la misma. Esta nueva pregunta —*cómo registrar lo que se observa*— alude ya a los diferentes *tipos de registros*. Así, si estamos haciendo auto-observación, podemos

registrar la información simplemente poniendo cruces en una lista de conductas ya preparada según lo que observamos, o anotando y describiendo todas las conductas que observamos, tal cual las observamos.

Aunque esta diferenciación entre modalidades y tipos de registros observacionales puede resultar más o menos clara en un principio, la cuestión es más compleja de lo que parece. Se puede hablar así de una *zona límite* en la que enfoques de observación y tipos de registros se entrecruzan, no resultando fácil distinguir con claridad entre todos ellos. Esto afecta especialmente a las distintas modalidades de observación, ya que los tres enfoques que hemos planteado (participación/no participación del observador; sistematización/no sistematización de la observación y auto/hetero-observación) no son independientes de los demás, de forma que cuando se realiza una observación hay que decidir en torno a los tres, y no en relación con uno de ellos de forma exclusiva. Las modalidades de observación serían, por tanto, tres continuos que adoptan una forma similar a un eje de coordenadas tridimensional (ver figura 1). En dicho eje, cada observación concreta puede ser situada en un punto determinado, según se aproxime más o menos, a cada uno de los seis polos de los tres continuos.

FIGURA 1
REPRESENTACIÓN DE LA RELACIÓN ENTRE LAS MODALIDADES DE OBSERVACIÓN



Quizá esta idea resulte mucho más clara con un ejemplo. Supongamos un profesor que observa las conductas de sus alumnos en la resolución de una tarea, utilizando para ello una lista en la que aparecen un conjunto determinado de conductas cuya ausencia-

presencia debe registrar. Si tuviéramos que situar esta observación mediante un punto en el eje de coordenadas de la figura 1, tendríamos que tener en cuenta que:

— El punto debe estar muy cercano al extremo de participación del eje participación/no participación, ya que el profesor no es un elemento extraño en la situación de aula.

— Al mismo tiempo, y dado que observa a otras personas y no a sí mismo, el punto debería estar también cercano al extremo de hetero-observación (en el eje auto/hetero-observación).

— Y, finalmente, puesto que ha definido con antelación qué conductas le interesa observar y decidido que sólo va a registrar su ausencia o presencia, el punto se aproximaría al extremo de sistematización (en el eje sistematización/no sistematización).

Obsérvese, por otra parte, que hablamos de «continuos» y sus extremos, y no tanto de dos opciones discretas. Esto nos permite indicar que en las tres modalidades de observación planteadas, existen multitud de grados de acercamiento a un polo u otro del continuo y no se trata de hacer una adscripción absoluta a una u otra forma de cada modalidad.

A continuación, profundizaremos en las distintas modalidades de observación enunciadas, haciendo una presentación inicial de los diferentes tipos de registros que se suelen utilizar en cada una de ellas.

1.2.1. Observación sistematizada/no sistematizada

La sistematización de la observación alude al hecho de si se determinan o no, con anterioridad al momento de la observación, los fenómenos a observar de forma precisa y controlada. En la medida en que las conductas a observar están previamente especificadas y definidas con precisión, estaremos ante un *procedimiento observacional sistematizado*. En el caso en que el observador sólo parta de algunas ideas vagas y generales de lo que quiere observar, estando este proceso abierto a las conductas reales que se manifiestan durante la observación, estaremos ante un *procedimiento observacional no sistematizado*.

ANGUERA (1988) clasifica las estrategias observacionales según estén más o menos sistematizadas. Distingue, bajo este criterio, tres diferentes maneras de enfocar la observación:

— *Observación no sistematizada*, ocasional o no controlada. Es aquella que no responde a una regla. Así, en la medida en que un problema de investigación es muy complejo y sus marcos teóricos de referencia están poco delimitados, es adecuado realizar observaciones no sistematizadas. Este tipo de observación se aplica cuando aún se sabe relativamente poco en torno al objeto a investigar. Se caracteriza por que la atención está poco estructurada pero abierta al máximo en

todas las direcciones y a todos los comportamientos que tengan lugar. De cualquier forma, cuando se emplea esta estrategia, el conocimiento que el observador tiene de la situación o conducta va aumentando, lo que conlleva a menudo que se vaya estrechando el campo de la observación.

— *Observación sistematizada* o controlada: de mayor precisión que la anterior, aunque restringe la libertad de observación a ciertos fenómenos. En ella, el observador prepara detalladamente, y con antelación al momento de la observación, el plan específico mediante el cual va a proceder.

— *Observación muy sistematizada*, que se diferencia de las anteriores por las siguientes propiedades: se construye sobre la base de una teoría explícita en función de la cual se aíslan conceptualmente las clases de variables de comportamiento que lo determinan; el proceso de observación se estructura en categorías definidas con precisión, habiendo de operacionalizarse las variables por medio de indicadores accesibles a la observación directa; las condiciones situacionales de partida son sometidas a control a fin de posibilitar observaciones comparables.

Sin embargo, otros autores prefieren usar los términos *abierto* o *cerrado* para designar una menor o mayor sistematización de la observación. Éste es el caso de EVERTSON y GREEN (1989). Los *sistemas cerrados* son los que contienen un número finito de unidades de observación prefijadas. Por lo general, las categorías que incluyen suelen ser mutuamente excluyentes y se las define de antemano de modo que reflejen determinadas convicciones teóricas respecto al fenómeno objeto de observación. Las observaciones se limitan a identificar y registrar las conductas contenidas en el propio sistema.

En los *sistemas abiertos*, siguiendo con EVERTSON y GREEN (1989), se utiliza una lente más amplia. Suelen tener una gama de categorías prefijadas que se utilizan para describir las conductas o pueden contener categorías generadas a partir de los patrones observados. En ellos, a menudo se registran conductas/fenómenos individuales mediante el uso de más de una categoría. Es posible combinar las categorías para reflejar la pluralidad de las manifestaciones conductuales. Otra característica de los sistemas abiertos es que pueden ser generativos, es decir, que hacen posible identificar nuevas variables durante todo el proceso de análisis.

Sintetizando estas aportaciones, las principales diferencias entre un sistema abierto y un sistema cerrado, pueden encontrarse en (EVERTSON y GREEN, 1989):

— El *punto temporal* en que se identifican las categorías. En los sistemas cerrados, las categorías se identifican a priori; en los sistemas más abiertos, normalmente se extraen categorías a partir de los datos obtenidos.

— El *rol del observador*. Mientras que en los sistemas cerrados, el observador se limita a registrar las conductas que figuran en el sistema; en los abiertos, el observador capta un segmento más amplio del contexto. Sus percepciones, formación y marco de referencia van a determinar lo que finalmente será registrado.

Esta diferenciación entre sistematización/no sistematización o, si se prefiere, apertura/no apertura de los sistemas de observación, no debe entenderse de forma absoluta, como ya se ha planteado con anterioridad. Así, caben establecer algunas diferencias entre los diversos procedimientos de registro utilizados tanto en la observación sistematizada como en la no sistematizada. Entendemos, por tanto, que el nivel de sistematización de un determinado tipo de registro observacional (o grado de control externo) es una dimensión que se extiende a lo largo de un continuo (ANGUERA, 1995b).

Por otra parte, existen diferentes tipos de registros a través de los cuales se pueden realizar observaciones sistematizadas y no sistematizadas. ANGUERA (1995b) hace una clasificación de los mismos, diferenciando entre:

1) Registros no sistemáticos: registro narrativo que puede presentar una de estas modalidades: diarios, registros anecdóticos, registros continuos y registros de muestras.

2) Registros con sistematización parcial: listas de control (con dos posibilidades: sistemas de signos y listas de rasgos) y escalas de estimación.

3) Registros sistematizados: datos categoriales.

También EVERTSON y GREEN (1989) clasifican diferentes registros abiertos y cerrados:

1) Registros categoriales. Generalmente están pensados para ser usados en un amplio espectro de situaciones, y no tanto en casos individuales. El interés suele ser identificar leyes generales de comportamiento.

2) Registros descriptivos. Generalmente se utilizan en combinación con registros tecnológicos (grabaciones en audio o vídeo). El período de observación puede variar en longitud, pero los datos se registran dentro de los límites naturales de los acontecimientos o contextos. Estos sistemas se dedican a obtener una descripción detallada de los fenómenos observados a fin de explicar procesos en desarrollo.

3) Registros narrativos. También suelen usarse en combinación con los sistemas tecnológicos. Los registros pueden realizarse *in situ* o en vivo (registro de incidentes críticos, registro de muestras y notas de campo) o pueden ser reflexivos (diarios). En ellos el período de observación puede variar según los propósitos del estudio. La duración de la observación puede consistir en un único acontecimiento o un período más prolongado. Aquí la preocupación también consiste en obtener descripciones detalladas de los fenómenos observados a fin de explicar procesos en desarrollo. El objetivo es no sólo comprender lo que está ocurriendo, sino también identificar factores que inciden en la aparición de esas conductas. En ellos, el observador es el principal instrumento de observación, ya que lo que se registra depende en gran medida de su sistema perceptual y de su capacidad para captar y transmitir en lenguaje cotidiano lo que ha observado. Al efectuar un registro narrativo, el observador registra un amplio segmento de la vida. No se busca filtrar

lo ocurrido de ningún modo sistemático. La información se recoge además desde una perspectiva cronológica.

4) Registros tecnológicos. Son grabaciones en vivo de acontecimientos y procesos. Proveen un registro permanente que suministra los datos en bruto sobre los que se debe trabajar para construir representaciones. Al tratarse de registros permanentes, es posible estudiar con mayor profundidad los acontecimientos en momentos posteriores al de la recogida de información. En ellos, es importante prestar atención a factores como el emplazamiento de la cámara o grabadora, el tipo de planos a tomar, la iluminación, etc.

En el apartado 3 de este capítulo profundizaremos en algunos de los registros mencionados, aportando algunas consideraciones para su diseño y aplicación en los procesos diagnósticos.

1.2.2. Observación participante/no participante

ANGUERA (1988) también aporta una clasificación de las observaciones de acuerdo con el criterio del grado de participación de la persona que realiza la observación. Diferencia así entre *observación externa o no participante, observación interna o participante* y *auto-observación*. Dejando este último tipo para un apartado posterior, a continuación iremos describiendo cada una de estas clases de observación.

La *observación externa o no participante* tiene lugar cuando el observador no pertenece al grupo objeto de estudio. En situaciones naturales, representa por sí misma una cierta artificialidad que «desnaturaliza» la observación. Así, el observador es un elemento extraño al contexto habitual y, por tanto, su rol puede modificar sensiblemente el proceso. La observación no participante, según ANGUERA (1988), puede ser directa o indirecta. La primera de ellas comprende todas las formas de observación hechas sobre el terreno en contacto inmediato con la realidad, aunque aquí estarían, en opinión de la autora, el cuestionario y la entrevista (técnicas que nosotros diferenciamos como procedimientos no observacionales). En realidad se trataría de observaciones que consisten en poner a prueba al sujeto ante una situación para recoger su comportamiento. La observación indirecta, a diferencia de la anterior, se basa en datos y fuentes documentales y su principal característica consiste en el hecho de que el investigador no ejerce control alguno sobre la forma en que los documentos han sido obtenidos, su actuación se limita a seleccionar, observar e interpretar lo que le interesa.

Las ventajas de este tipo de observación radican en que el observador puede dedicar toda su atención a la observación y a realizar las anotaciones al mismo tiempo que se originan los fenómenos. El inconveniente radica en las repercusiones que puede conllevar la presencia de un observador (ANGUERA, 1988).

Por su parte, la *observación participante (o interna)*, según ANGUERA (1988), persigue recoger la información de la conducta a través del contacto directo con la realidad en que ésta tiene lugar. El observador que recoge la información, registra e interpreta los datos

al participar en la vida diaria del grupo u organización que estudia, entrando en conversación con sus miembros y estableciendo alguna forma de asociación o estrecho contacto con ellos. El rasgo básico en las diferentes modalidades de observación participante es, según esta autora, que la persona que realiza la observación se gane la confianza de las personas a las que observa, de tal forma que su presencia no perturbe o interfiera en el curso natural de los acontecimientos. Aquí va a ser realmente importante la capacidad y formación del observador para integrarse en el entorno en el que observa y compartir sus modos de pensar y vivir. Su éxito radicará en que sea capaz de observar sin los prejuicios y preconceptos propios.

Este tipo de observación presenta la ventaja de permitir al observador captar la perspectiva interna de los acontecimientos y registrarlos tal como fueron percibidos por alguien que participó en ellos (EVERTSON y GREEN, 1989). Aunque también se le critica la subjetividad a la que puede conducir el hecho de que el observador sea partícipe de la realidad y situación que observa.

1.2.3. Auto-observación/hetero-observación

Cuando se realiza una observación es habitual que ésta recaiga sobre otras personas distintas al observador: una situación de clase, los comportamientos de una familia, una reunión de docentes... Este tipo de observaciones, habituales y frecuentes en los procesos diagnósticos, pueden considerarse como *hetero-observación*, en tanto la persona que observa es distinta a la/s observada/s.

Sin embargo, también existen muchas situaciones en las que, con una finalidad diagnóstica, el observador se convierte en el propio objeto de la observación. Estaríamos, pues, ante una *auto-observación*. La auto-observación ha sido definida como un doble proceso consistente, por un lado, en atender deliberadamente a la propia conducta y, por otro, en registrarla por algún procedimiento previamente establecido (AVIA, 1983).

Esta modalidad de la observación tiene la peculiaridad de servir tanto a funciones evaluativas y diagnósticas, como a funciones terapéuticas o intervencionistas (HERBERT y NELSON-GRAY, 1997). En el primer caso, función evaluativa, nos resulta útil para acceder a determinados fenómenos, situaciones y conductas que tienen lugar de forma privada (por ejemplo, las que ocurren dentro del marco familiar, que no siempre son accesibles al profesional del diagnóstico) o de manera íntima (lo que sucede en el interior de la persona y que no necesariamente se manifiesta siempre en comportamientos observables por otra persona).

En su vertiente terapéutica, es evidente que cuando una persona se observa a sí misma está desarrollando su capacidad de introspección o su atención consciente a determinados eventos que se producen en su ambiente y que tienen una repercusión en ella. Es, por tanto, una forma de empezar la intervención orientadora y educativa, en tanto que proporciona un mecanismo para reforzar la sensibilidad del sujeto y su toma de

conciencia sobre los aspectos que condicionan un problema o necesidad educativa. Al mismo tiempo, la técnica de la auto-observación tiene también la ventaja de fomentar la motivación y la implicación de la persona objeto de diagnóstico. HERBERT y NELSON-GRAY (1997) sugieren que el hecho mismo de autorregistrar una conducta hace a menudo que cambie la frecuencia de dicha conducta. Esta reactividad demuestra ser terapéutica en tanto los cambios que tienen lugar en la conducta se manifiestan en la dirección adecuada. Así, las conductas evaluadas positivamente tienden a aumentar, mientras que las consideradas negativas tienden a disminuir.

Si bien la auto-observación no está libre de cierta subjetividad a la hora de seleccionar y registrar la información, no cabe duda de que es una estrategia que aporta considerables ventajas en los procesos diagnósticos.

2. PREPARACIÓN DE LA OBSERVACIÓN: LA SELECCIÓN DE MUESTRAS

Dado que los fenómenos que observamos ocurren en un tiempo y situación determinados, y con unos protagonistas concretos, además de la necesaria pregunta acerca de *qué* observar, es fundamental que también nos planteemos *cuándo* y *cómo* hacerlo, siendo igualmente conveniente, en algunos casos, decidir en torno a *quién* observar. Todas estas decisiones conforman lo que se denomina la *selección de muestras* de observación, y tienen una importancia crucial en los procesos observacionales porque, en tanto se resuelvan adecuadamente, van a garantizar la significatividad y representatividad de nuestras observaciones.

Es evidente que la observación se encuentra sujeta a múltiples limitaciones. Muchas conductas tienen lugar en ámbitos que no son accesibles de forma directa al observador (por ejemplo, el contexto familiar); otras son de carácter privado e íntimo (por ejemplo, la conducta sexual). En ocasiones, deseamos observar el comportamiento de muchos sujetos, lo que implica la imposibilidad de centrar la atención en más de uno a la vez. Incluso si todas estas limitaciones no existieran, la capacidad de atención del ser humano no puede prolongarse indefinidamente; el cansancio y la fatiga provocarían una importante pérdida de efectividad en el registro.

Ante la existencia de estas limitaciones que nos impiden observar todo, nos vemos obligados a seleccionar «muestras» para nuestra observación. En la medida en que la selección de estas muestras se realice conforme a ciertos criterios, conseguiremos que las mismas puedan considerarse como representativas de los fenómenos a observar.

No obstante, la selección de muestras en la observación no se refiere de forma exclusiva a los sujetos que van a ser observados, sino que incluye las decisiones relativas al tiempo y a la situación en que se va a observar. FERNÁNDEZ BALLESTEROS (1992) resume estas decisiones en el siguiente conjunto de preguntas:

1. Durante cuánto tiempo va a prolongarse la observación.
2. Con qué frecuencia se va a observar.
3. En qué momentos se van a iniciar y terminar los períodos de observación y si éstos van a ser constantes o variables.
4. Si se van a utilizar o no intervalos de tiempo para la observación seguidos de intervalos para el registro.
5. En qué situaciones se va a observar.
6. Si se va a observar a un sujeto o a varios.

Todas estas decisiones aluden a diferentes tipos de muestreo que se pueden clasificar en tres grupos: muestreo de tiempo, de situaciones y de sujetos (IBÁÑEZ, 1993). Aunque a continuación profundizaremos en cada uno de ellos, conviene saber que, según la forma en que enfoquemos la observación y el tipo de registro que utilicemos, no siempre estas preguntas tendrán el mismo grado de importancia, además de que la respuesta a las mismas puede ser muy diferente. Así, cuando hacemos una observación no sistematizada y participante sobre una situación, probablemente seleccionemos el marco temporal, espacial y personal de la observación de forma muy distinta a como lo haríamos si fijamos un conjunto de conductas previamente y observamos su ocurrencia en una realidad de la que no somos partícipes.

En los registros más abiertos y participantes, la observación suele acogerse a los *límites naturales* dentro de los cuales se manifiestan las conductas o fenómenos objeto de estudio. Tomemos como ejemplo una situación de enseñanza-aprendizaje que se desarrolla dentro de un aula; probablemente se seleccione como período de observación la clase completa, y no fragmentos temporales discontinuos e inconexos. La duración de la observación también tendrá como referente las unidades temporales importantes para el objeto de estudio. En nuestro ejemplo, la duración de la observación podría abarcar un curso escolar, un trimestre... Asimismo, la situación en la que se observa será el marco natural en que se desarrolla el acontecimiento o, dicho de otra forma, el lugar o lugares donde ocurre aquello que nos interesa observar.

Hechas estas aclaraciones, pasamos ya a explicar los diferentes tipos de muestreo, comenzando por el **muestreo de tiempo** que es quizá el que más decisiones conlleva ya que, dentro de él, si recordamos la cita presentada de FERNÁNDEZ BALLESTEROS (1992), se engloban las cuestiones relativas a cuánto debe durar la observación, con qué frecuencia va a tener lugar ésta, en qué momento se va a observar y si los períodos de observación van a ser divididos o no en intervalos. Todas estas decisiones van a depender, en primer lugar, de qué se está observando, y como planteábamos anteriormente, del enfoque con que se observa.

No existen prescripciones concretas en torno al tiempo que debe durar una observación. Por un lado, cuanto más tiempo se observe, mayores posibilidades habrá de generalizar las conductas o fenómenos observados. Sin embargo, hay que tener en cuenta el coste temporal, económico y de esfuerzo que supone prolongar en exceso una observación. Por eso, en primer lugar, la duración de la observación va a depender de las características del fenómeno que se observa. Entre otras cosas, se aconseja que la duración de la observación sea inversamente proporcional a la frecuencia del evento observado (FERNÁNDEZ BALLESTEROS, 1992). Así, si una conducta tiene lugar con frecuencia, es posible que en poco tiempo hayamos tenido la posibilidad de observarla, por lo que no es necesario prolongar en el tiempo la observación. Sin embargo, si el evento o conducta es infrecuente o raro, resultará difícil apreciarlo si el tiempo total de duración de la observación es demasiado corto.

Muy en relación con la duración de la observación, está la cuestión de la frecuencia con la que se observa. Así, ante una observación que se prolonga durante un trimestre, no es lo mismo observar una vez al día, que una vez a la semana o que una vez a la quincena.

La elección de los momentos en que comenzarán y finalizarán las sesiones de observación o muestreo intersesional, según QUERA (1991), es también necesario, no sólo porque no podemos pasarnos todo el día observando, sino también porque, aunque pudiéramos, la calidad de nuestras observaciones descendería considerablemente, al entrar en juego la fatiga y el cansancio que provocaría sostener de forma constante nuestra atención. En este sentido, las sesiones deben cubrir de la forma más representativa posible todo el período de estudio, y en la medida de lo posible, no depender de las preferencias o disposiciones temporales del observador. Ello implica que los momentos de inicio de las sesiones han de ser seleccionados de manera que todos los puntos temporales del período objeto de estudio tengan la misma probabilidad de ser inicios de sesiones.

La selección de los períodos de observación, no obstante, dependerá en primer lugar del fenómeno, evento o conducta que estamos estudiando. Así, si queremos observar los comportamientos no verbales en el aula de los estudiantes de secundaria, nuestros períodos de observación deberán regirse por el horario escolar y abarcar todo o parte del tiempo de duración de una sesión de clase. De cualquier forma, los períodos pueden ser constantes, en cuyo caso tienen lugar a la misma hora todas las sesiones de observación (por ejemplo, todas las tardes de 4 a 5) o, variables, si se va alternando la hora con algún criterio (por ejemplo, de 4 a 5 la primera sesión, de 5 a 6 la segunda, de 6 a 7 la tercera, etc.).

Una de las cuestiones más complejas, y a la vez más importantes, en la selección de muestras temporales es cómo va a ser considerado el período de observación. Así, la sesión de observación puede convertirse en un único intervalo en el que se produce un registro continuo y completo de todos los eventos o conductas. Sin embargo, esto exige una atención constante por parte de la persona que realiza la observación, por lo que también se puede llevar a cabo un registro discontinuo, en el que se establece una longitud de intervalo temporal con el que se dividirá la sesión de observación, y que se utilizará como pauta para el registro de la conducta (CARRERAS, 1991).

Por lo tanto, existen dos criterios posibles a la hora de observar (QUERA, 1991): por un lado, anotar las ocurrencias de las categorías en la secuencia en que se producen, registrando en qué momento acaba una y en qué momento comienza la siguiente; esto es lo que se llama un *registro activado por transiciones* (RAT), en el que el observador únicamente registra los momentos en que se produce una transición entre ocurrencias de categorías. Por otro, podemos hacer registros tomando como referencia unidades de tiempo (intervalos). Esto es lo que se llama un *registro activado por unidades de tiempo*

(RAUT), en el que es la finalización de cada unidad temporal lo que activa al observador para que anote.

Existen diferentes modalidades de RAUT (QUERA, 1991; QUERA y BEHAR, 1997):

— *Intervalo instantáneo, momentáneo o puntual*: la sesión de observación es dividida en N intervalos y se registra sólo la categoría que tiene lugar en un momento determinado del intervalo (al principio, al final...).

— *Intervalo parcial*: la sesión se divide también en N intervalos y el observador registra en cada uno de ellos qué categorías han ocurrido (en algunos casos también cuál ha sido su duración y en qué momento del intervalo).

— *Intervalo total*: la sesión se divide en N intervalos y se anota en cada uno de ellos cuál ha sido la categoría que ha ocupado la totalidad del intervalo (así, en los intervalos en los que acontece más de una categoría no hay registro).

Según IBÁÑEZ (1993), la duración del intervalo depende en principio del tamaño de las conductas. Así, conductas de escasa duración requieren intervalos mayores que conductas de más duración. En general, se consideran adecuados intervalos de entre 5 y 30 segundos, aunque autores como BAKEMAN y GOTTMAN (1989) recomiendan no sobrepasar los 10-15 segundos. El intervalo establecido no debe ser tan largo que mezcle inicios y finales de los eventos que se están estudiando.

Asimismo, en los muestreos de intervalo parcial y total, es posible que el observador pierda detalles conforme fija su atención en el papel para registrar las categorías. Por ello, se puede establecer un sub-intervalo de registro después del intervalo de observación. Además, cuando los períodos de observación son especialmente largos, pueden establecerse intervalos de descanso.

En el cuadro 1 se muestra una posible organización de la sesión de observación. La sesión dura aproximadamente media hora, en la que se diferencian un total de 80 intervalos de observación y registro, existiendo un intervalo de descanso justo en la mitad de la sesión. El tiempo total dedicado a la observación —es decir, sin contabilizar el tiempo dedicado al registro o al descanso— es de 20 minutos. Las conductas observadas y registradas para este período de 20 minutos serán generalizadas al total del tiempo observado (los 27 minutos que dura la sesión).

CUADRO 1

EJEMPLO DE SECUENCIA DE LOS INTERVALOS DE OBSERVACIÓN, REGISTRO Y DESCANSO

Comienzo: 8:00h								
Obs. 1	Reg. 1	Obs. 2	Reg. 2	Obs. 3	Reg. 3 (...)	Obs. 40	Reg. 40	Descanso (8:13 h)
15"	5"	15"	5"	15"	5"	15"	5"	40"
Obs. 41	Reg. 41	Obs. 42	Reg. 42	Obs. 43	Reg. 43 (...)	Obs.80	Reg.80	Fin (8:27h)
15"	5"	15"	5"	15"	5"	15"	5"	

El **muestreo de sujetos** guarda cierta relación con el muestreo de tiempo. Aunque, primero, hemos de diferenciar lo que sería la selección de la muestra (o la decisión relativa a qué personas vamos a observar), de lo que sería la selección de sujetos cuando observamos dentro de un grupo. QUERA (1991) denomina muestreo de sujetos a la selección de la muestra, mientras que prefiere hablar de *muestreo intrasesional* de sujetos para referirse a la selección de sujetos en un grupo. Así, es evidente que cuando se está registrando el comportamiento de un grupo, es imposible que el observador mantenga simultáneamente la atención sobre todos sus componentes (dando por supuesto que todos los sujetos del grupo forman parte de la muestra). La selección del sujeto que se va a observar en cada momento es, por tanto, el muestreo intrasesional. Generalmente se combina con el muestreo de tiempos, de tal forma que se observa a un sujeto en un intervalo o en una sesión y a otro sujeto en otro intervalo u otra sesión. Cuando en cada sesión se observa a un solo sujeto se utiliza la estrategia denominada *muestreo focal*; si, dentro de una misma sesión, en cada intervalo (y por turnos) observamos a un sujeto, utilizamos la estrategia de *muestreo multifocal o muestreo de barrido*. Aquí el orden en que se observa a cada sujeto es siempre el mismo en cada turno y se establece de antemano. Además, el tiempo que observamos a cada sujeto es considerablemente menor que en el muestreo focal, aunque, obviamente, lo observamos más veces.

Sea cual sea el tipo de muestreo intrasesional de sujetos por el que nos decantemos, lo importante, nuevamente, es que los resultados procedentes de los sujetos observados puedan generalizarse al conjunto de sujetos que componen la población objeto de diagnóstico.

Pasamos ya a hablar del **muestreo de situaciones.** Obvia decir que debemos observar en todas aquellas situaciones en las que la conducta que nos interesa pueda ocurrir, a fin de que podamos obtener una muestra representativa de dichas situaciones. Asimismo, en ocasiones, resulta necesario conocer si un determinado comportamiento es estable y aparece en situaciones distintas o si, por el contrario, depende de una o varias situaciones concretas.

Conviene, por tanto, establecer un listado de todas las situaciones posibles en las que interesa observar y tomar muestras de la conducta en todas ellas. Como señala FERNÁNDEZ BALLESTEROS (1992), el muestreo de situaciones ha de realizarse en función de los objetivos del diagnóstico y, por tanto, basarse en criterios racionales.

Relacionada con el muestreo de situaciones, pero abarcando un campo mucho más amplio, está la cuestión del **contexto de la observación,** siendo éste un factor muy importante a tener en cuenta a la hora de poner en práctica procedimientos observacionales. Sin embargo, como plantean EVERTSON y GREEN (1989), existen diferentes formas de considerar el contexto:

— El *contexto local inserto en niveles de contexto más amplios*. Además del contexto inmediato al acontecimiento observado, existen otros contextos más amplios que influyen en el primero. Así, un grupo de alumnos está inserto en el contexto aula, ésta a su vez, está inserta en la escuela, ésta a su vez en un barrio o zona... Entre todos estos contextos existen vínculos que se reflejan en el contexto inmediato de observación, por lo que es necesario considerar múltiples niveles de contexto.

— El *contexto histórico del lugar.* Que también puede ayudar a la hora de explicar las situaciones y conductas presentes.

— El *contexto histórico del acontecimiento*. Los procesos y conductas se desarrollan a lo largo del tiempo, y esta perspectiva histórica es importante a la hora de entender los acontecimientos que se observan.

Además de todas las dimensiones contextuales descritas, EVERTSON y GREEN (1989) también se refieren al contexto según lo determina el enfoque de la observación. Con esto se refieren a la perspectiva respecto al contexto que se adopta al hacer la observación. Así, si la recopilación de datos tiene en cuenta los aspectos contextuales, estaremos ante un enfoque *inclusivo*, mientras que si se obvia la referencia contextual, el enfoque es *exclusivo*. Este último enfoque tiende a centrarse en las conductas minimizando lo que ocurre en el contexto. En el extremo de los enfoques inclusivos no se hace ningún intento por excluir los aspectos contextuales, sino que se procura abarcar el fragmento más amplio posible de la realidad observada. Ambos enfoques no tienen por qué ser mutuamente excluyentes y se pueden aplicar de forma conjunta. Esto es lo que EVERTSON y GREEN (1989) denominan como un *cambio de lente*, en tanto que con el uso de distintos procedimientos y enfoques se puede profundizar más en la realidad y en sus distintas dimensiones.

En definitiva, a la hora de realizar observaciones, además de prestar atención al procedimiento o procedimientos que vamos a utilizar, es necesario plantearse un conjunto de cuestiones relativas a la planificación y preparación de la observación. ¿Cuánto va a durar la observación?, ¿con qué frecuencia se va a observar?, ¿a qué sujetos se observará?, ¿en qué situaciones?, ¿qué contextos se van a considerar?... En este apartado, hemos intentado ofrecer algunos criterios para responder a estas preguntas,

dejando claro, no obstante, que la modalidad de observación y el tipo de registro condicionarán en gran medida la toma de decisiones.

3. DISEÑO Y APLICACIÓN DE PROCEDIMIENTOS OBSERVACIONALES

En el apartado 1 hemos presentado diversas formas de enfocar la observación y diferentes registros asociados a cada una de ellas. Se trata ahora de profundizar algo más en cómo se diseñan y aplican algunos de los principales registros observacionales. Para ello nos encontramos con una primera dificultad que se refiere a qué tipo de registros considerar, ya que no existe un consenso generalizado entre los diversos autores que establezca de forma clara cuáles son los mismos y en qué se diferencian unos de otros.

Para complicar aún más este panorama, algunos autores (ANGUERA 1995c; SIERRA BRAVO, 1983) parten de un concepto más amplio de la observación, incluyendo dentro de los registros observacionales la entrevista y el cuestionario, técnicas que nosotros hemos clasificado dentro de los métodos de encuesta y que serán abordadas en el siguiente capítulo. Asimismo, otras técnicas de autoinforme como las historias de vida y las autobiografías también son incluidas en ocasiones (BEHAR, 1991) dentro de las estrategias observacionales, en tanto que implican la puesta en práctica de procesos de auto-observación.

En nuestro caso, vamos a considerar dos grandes grupos de registros observacionales: por un lado, los *registros sistematizados*, entre los que hablaremos de los sistemas categoriales, las listas de control y las escalas de estimación, por ser los más ampliamente utilizados en el campo educativo; por otro, los *registros no sistematizados*, dentro de los cuales hablaremos de las notas de campo, el diario y el registro de anécdotas, también en este caso, los más usados en la práctica del diagnóstico en educación. Ya expusimos anteriormente que la sistematización de la observación es una cuestión de grados dentro de un continuo, por lo que entendemos que, si bien con nuestra clasificación no recogemos la totalidad de registros observacionales existentes, al menos, llamamos la atención sobre los más conocidos que, a su vez, representan alternativas diversas a la hora de observar.

Por otra parte, queremos aclarar que, aunque hemos tomado como referencia, a la hora de hacer nuestra clasificación, el grado de sistematización de la observación, los diferentes tipos de registros que se van a presentar pueden adquirir distintos formatos según las otras dos modalidades de observación que distinguíamos con anterioridad. Así, por poner un ejemplo, a la hora de hacer auto-observación se pueden utilizar tanto procedimientos abiertos —como el diario—, como procedimientos cerrados —como una escala de estimación—. Lo que variará en cada caso es la forma en que se enfoca la observación —centrada en la propia persona— y la forma en que se usa y aplica el registro. A pesar de ello, hay ciertos registros que son difíciles de compaginar con alguna/s modalidad/es de observación. El ejemplo más claro que se nos ocurre es usar un sistema categorial dentro de una observación participante. Como veremos a

continuación, los sistemas categoriales exigen una alta concentración —generalmente hay que registrar conductas de acuerdo con determinados intervalos temporales, lo que implica estar muy pendientes del cronómetro además de las conductas a observar—, que resulta difícil de compaginar con la participación del observador en los fenómenos que observa. Por ello, puede decirse que este tipo de registros sólo permiten un nivel medio de participación (por ejemplo, un profesor que observa a sus alumnos pero sin participar en lo que ellos hacen).

Por último, queremos hacer una breve mención al *análisis de documentos*, que también suele ser considerado en algunas clasificaciones (ANGUERA, 1995c; BEHAR, 1991; SIERRA BRAVO, 1983) como un procedimiento observacional. Cuando se analiza el contenido de material documental como cartas, documentos oficiales (por ejemplo, en un centro educativo, sus diferentes planificaciones y proyectos), libros de texto, material didáctico, etc., se suelen utilizar dos formas genéricas de proceder:

— Por un lado, determinando a priori un conjunto de características que se consideran relevantes que estén incluidas o no en el documento, y registrando su presencia/ausencia o la intensidad o frecuencia con que aparecen.

— Por otro, leyendo detenidamente el material y extrayendo de él los aspectos que se consideren más relevantes y significativos.

En el primer caso, pueden utilizarse cualquiera de los procedimientos que vamos a presentar dentro de los registros no sistematizados, especialmente las listas de control y las escalas de estimación. En su elaboración, se habrá de tener presente que no se trata de observar personas y sus interacciones, pero de cualquier forma, el procedimiento puede ser muy similar.

En el segundo caso, en realidad estaríamos procediendo de acuerdo con lo que habitualmente se denomina como *análisis de contenido* (Bardin, 1986), estrategia que se utiliza para analizar, entre otras cosas, los registros no sistematizados como el diario o las notas de campo. Luego, de alguna forma, la observación de material documental puede considerarse incluida, de forma genérica, dentro de la clasificación que aquí estamos utilizando.

Presentamos ya los principales registros observacionales, recordando una frase de EVERTSON y GREEN (1989:320) que ofrece un consejo para su selección:

> «Lo fundamental no es qué instrumento es mejor, sino cuál es el más apropiado para la cuestión que se estudia y cuál representará adecuadamente el segmento de la realidad en observación».

3.1. Observación sistematizada

Una forma de evitar algunos de los errores e imprecisiones de la observación abierta (prejuicios del observador, percepción selectiva, vaguedad de sus observaciones...) es

usar normas que sistematicen la observación. En este caso, estaríamos poniendo en práctica una observación sistematizada. Entre las características atribuidas por ANGUERA (1988) a este tipo de observación, hemos seleccionado las siguientes:

— Es específica y cuidadosamente definida, lo que permite que el observador pueda determinar con antelación las categorías a observar.

— El tipo de datos a recoger también se fija previamente.

— Generalmente es cuantificable.

Se trata, por tanto, de una modalidad de observación en la que se delimitan y especifican, con anterioridad a la observación, las conductas o eventos que van a ser observados. El acto de observar en sí consiste básicamente en prestar atención a la ocurrencia de las conductas previamente definidas, registrando su aparición, frecuencia, duración, intensidad... y obviando cualquier otro evento que no haya sido previamente considerado. Esta forma de proceder, además de facilitar considerablemente la observación (ya que el observador sabe en todo momento qué tiene que observar y registrar), permite que distintos observadores se fijen siempre en las mismas conductas y, por tanto, sus observaciones sean directamente comparables, además de cuantificables.

Ahora bien, existen diferentes tipos de registros a través de los cuales se realizan observaciones sistematizadas: sistemas de categorías, listas de control, escalas de estimación, sistemas de signos, etc. Cada uno de ellos presenta unas características diferenciales, si bien por tratarse de registros sistematizados comparten unos rasgos en común. En este apartado haremos mención de tres de estos sistemas, por ser los más ampliamente utilizados en educación: los sistemas de categorías, las listas de control y las escalas de estimación. EVERTSON y GREEN (1989) los incluyen dentro de un mismo grupo genérico al que denominan *sistemas categoriales*. Dadas las similitudes entre estos sistemas, en este apartado, vamos a presentar de forma conjunta el proceso a seguir para su construcción, para posteriormente hacer una breve presentación de cada uno de ellos por separado.

3.1.1. Aspectos a tener en cuenta en la elaboración de sistemas categoriales

ANGUERA (1991:116) describe el sistema de categorías como «la construcción de una especie de andamiaje que proporcione soporte y cobertura a aquellas conductas que, mediante la correspondiente operación de filtrado, son consideradas relevantes de acuerdo con los objetivos de la investigación, y todo ello con un máximo de flexibilidad que posibilite la adaptación al flujo de conducta tal cual transcurre y a la situación y contexto en que se inscriba».

El desarrollo de un sistema de categorías es, por tanto, un acto de naturaleza teórica (ANGUERA, 1991), y su elaboración y aplicación no tienen sentido si no existe un marco teórico previo desde el cual se seleccionan y definen las conductas relevantes que van a

ser observadas. Recordemos que estos tipos de registros han sido calificados anteriormente como «sistematizados» o «cerrados», por lo que su característica principal es esa sistematización o estructura que se ha elaborado en torno a la pregunta o problema que suscita la observación.

En esta línea se manifiestan también BAKEMAN y GOTTMAN (1989), para quienes formular una pregunta clara a la que dar respuesta es el primer paso necesario para elaborar un sistema categorial. A juicio de estos autores, el desarrollo de un sistema de codificación en el que se seleccionan un conjunto de conductas a observar es un trabajo teórico y no tanto mecánico o técnico y, como tal, debe enfocarse.

Si desconocemos el tema sobre el cual va a girar la observación, si no sabemos *a priori* cuáles son las conductas que pueden resultar más relevantes, si nos resulta difícil distinguir entre algunos comportamientos, etc., es mejor que optemos por un sistema abierto para nuestra observación. De esta forma, podremos explorar sobre el terreno los eventos y conductas sin cerrar el alcance de nuestra observación a comportamientos o situaciones que pueden ser poco relevantes.

De hecho, con anterioridad a la elaboración de un sistema categorial, además de las necesarias lecturas que facilitarán la construcción de un marco teórico-conceptual sobre el cual construir las categorías, es recomendable hacer previamente una observación exploratoria, más informal y menos sistematizada, que nos permita conocer mejor la situación o fenómeno a observar. Esto nos ayuda no sólo a conocer cuáles son las conductas más relevantes, frecuentes y típicas, sino que también, de forma indirecta, contribuirá a definir mejor nuestro objetivo diagnóstico.

Una vez definido el marco teórico que sustenta la observación, será posible formular las categorías en torno a las cuales se realizará dicha observación. El concepto de categoría es clave en los sistemas categoriales, en tanto que éstas son las unidades de conducta que orientan la observación. Una categoría es (ANGUERA, 1991:120):

> « el resultado de una serie de operaciones cognitivas que llevan al establecimiento de clases entre las cuales existen unas relaciones de complementariedad, establecidas de acuerdo con un criterio fijado al efecto, y donde cada una de ellas cumple a su vez requisitos internos de equivalencia en atributos esenciales, aunque pueda mostrar una gama diferencial o heterogeneidad en su forma».

Así, si los sistemas categoriales son construcciones conceptuales en las que se operativizan las distintas manifestaciones del fenómeno objeto de observación, la categoría no es sino una clase dada de ese fenómeno, según una regla de correspondencia unívoca. En el apartado siguiente, veremos, no obstante, algunas normas que toda categoría debe cumplir.

Existe un conjunto de cuestiones que afectan a la construcción de un sistema categorial, y de cuya resolución depende el producto final de este proceso de elaboración, que será el instrumento diseñado (el sistema de categorías, lista de control o escala de estimación que finalmente utilicemos para observar la realidad). Todas estas

cuestiones conllevan un conjunto de decisiones ante las que tiene que optar el constructor del instrumento. A continuación, veremos las más importantes de ellas.

A la hora de generar un sistema categorial podemos actuar desde diferentes niveles de descripción. Así, podemos hacer referencia a fenómenos directamente observables (levanta la mano, tira el lápiz...) o a otros que requieren un mayor *nivel de inferencia* (se pone triste, está implicado en la tarea...). Por tanto, en los sistemas en los que el nivel de inferencia es bajo, se registran las conductas tal cual suceden (generalmente, se codifican), mientras que en aquellos en los que el nivel de inferencia es alto, el observador se ve obligado a razonar o interpretar la conducta (por ejemplo, registrar acciones como «se siente contento en el grupo», que no se observan directamente, sino que se infieren a partir de un cúmulo de comportamientos observables).

La cuestión de la inferencia que soporta un sistema categorial nos remite a la elección de *conductas molares* (amplias y generales) *o moleculares* (específicas y concretas). BLANCO y ANGUERA (1991) caracterizan las unidades de observación como molares si son integradoras y abarcan diferentes respuestas del sujeto de forma global, llegando a ellas a través de la abstracción; mientras que las moleculares ofrecen una terminología básica de la conducta observada, refiriéndose simplemente a la descripción de atributos observables.

Así, generalmente las conductas molares poseen mayor significado pero requieren un mayor grado de inferencia y, por tanto, resulta más difícil demostrar su validez. Sin embargo, las categorías moleculares son más fáciles de observar, ya que son conductas definidas más discretamente, aunque ello puede suponer una mayor dificultad en cuanto a su interpretación y dotación de significado (por ejemplo, resulta difícil decidir si golpear la mesa con fuerza denota o no un comportamiento agresivo).

En relación con el problema de la inferencia en los sistemas categoriales, BLANCO y ANGUERA (1991) plantean que la inferencia ha de ser situada en un continuo. Así, en una observación con una débil inferencia, el observador enuncia escrupulosamente aquello que ve o entiende sin preocuparse del significado que los hechos conllevan. En una observación con una fuerte inferencia, el observador puede enunciar intenciones, motivos, sentimientos... es decir, atribuye significados a los fenómenos que observa.

En nuestra opinión, cuando se construye un sistema categorial, debe optarse por categorías moleculares, directamente observables. En el momento en que las categorías soportan un gran nivel de inferencia, es habitual que resulten poco claras y no cumplan el requisito de mutua exclusividad que entre ellas debe darse. Los sistemas de categorías que están basados en conductas molares, si bien pueden ser válidos, requieren un esfuerzo teórico mucho mayor a fin de evitar los problemas descritos, y sólo en pocos casos esto se consigue. Si se prefiere una observación que registre los significados y las inferencias, es mejor optar por otras modalidades y registros a la hora de observar, como las que se presentarán en el apartado relativo a la observación no sistematizada. Mientras

que cuando se opta por un sistema categorial, se ha de procurar que las categorías respondan a los requisitos científicos y técnicos que habitualmente se les prescriben.

BAKEMAN y GOTTMAN (1989) también se muestran partidarios de la molecularización, por varias razones: a) se incrementa la probabilidad de obtener datos fiables; b) siempre se pueden agrupar *a posteriori* (cosa que no es posible en sentido contrario); c) las categorías moleculares, por su menor grado de inferencia y mayor objetividad, pueden revelar datos de interés que ayuden a justificar los resultados (esto no es factible con las molares por ser más difusas y susceptibles de dar lugar a interpretaciones heterogéneas).

Otra cuestión muy importante en los sistemas categoriales se refiere a cuántas *dimensiones* los componen. Así, existen sistemas unidimensionales que observan comportamientos homogéneos que se refieren a una misma esfera de características, mientras que otros describen un determinado comportamiento a partir de un cierto número de categorías. Aunque la mayoría de los sistemas de categorías suelen tener varias dimensiones, y esto implica que abarcan más facetas de la realidad objeto de estudio, hay que tener en cuenta que cuantas más dimensiones hay en un sistema, se producirá un menor nivel de acuerdo entre los observadores (ANGUERA, 1988).

Por otra parte, las categorías de un sistema pueden ser *continuas o discontinuas* (discretas). Hablamos de categorías continuas cuando las conductas que recogen se encuentran situadas en un continuo. Siguiendo el ejemplo propuesto por ANGUERA (1988), supongamos que se establece un cierto número de categorías que permiten clasificar las diversas manifestaciones agresivas como: I) ataques verbales moderados; II) amenazas verbales y gestuales; III) agresión física; etc. Todas estas conductas implican una cierta gradación. Así, la de menor intensidad sería la I, y la de mayor la III. En este sistema, es fácil que todos estemos de acuerdo con que agredir físicamente es una conducta violenta más intensa que amenazar. Sin embargo, no todos los fenómenos admiten teóricamente que sus manifestaciones conductuales puedan ser situadas en un continuo. En esos casos, estamos ante sistemas con categorías discretas o discontinuas, que suelen ser los más habituales en el campo educativo.

Los sistemas de categoría también pueden diferir según la *unidad de conducta* que deberá clasificarse y registrarse. Es decir, a la hora de construir un sistema categorial, es necesario responder a la pregunta «¿qué cantidad de conducta es clasificable en cada categoría?». Generalmente para responder a esta pregunta, se suele utilizar el tiempo como criterio. Es decir, se trata de ver cuál es la duración o la frecuencia de la conducta. Para ello, y como se ha visto en apartados anteriores, se determinan intervalos de observación. En este sentido, conviene saber que cuanto más cortos son los intervalos de tiempo, mayor es la fiabilidad, y los errores tienden a ser bajos porque los observadores adoptan fácilmente un ritmo y son más constantes en sus anotaciones (ANGUERA, 1988).

Por último, queremos hacer referencia a otra cuestión que, según KERLINGER (1975), atañe a la construcción y aplicación de sistemas categoriales: la que se refiere a su *generalidad*, o grado de aplicabilidad a situaciones diferentes. Algunos sistemas tienen

una aplicación más restringida, en tanto que sólo pueden ser aplicados a situaciones concretas y determinadas, mientras que otros recogen conductas a un nivel de generalidad que permite su aplicación a un conjunto amplio de situaciones.

BAKEMAN y GOTTMAN (1989) y ANGUERA (1991) se muestran partidarios de que los sistemas categoriales se elaboren *ad hoc* para conseguir una mayor adaptación a la situación o comportamiento a observar. Si bien es evidente que esto tiene sus ventajas, no debe olvidarse que la construcción de sistemas categoriales requiere una gran inversión y dedicación por parte de la persona que hace el diagnóstico. En su grado máximo, algunos sistemas de categorías requieren procesos y tareas similares a las necesarias para construir los tests. Es por ello que pensamos que estos sistemas deben ser «rentables», en tanto sean versátiles y susceptibles de ser aplicados a muchas situaciones distintas. Por poner un ejemplo, un sistema categorial que pueda aplicarse en cualquier situación de aula siempre será más generalizable que aquél que sólo es válido para situaciones de aula en la que se trabajan problemas matemáticos. Por lo tanto, sin perder de vista los fines del diagnóstico en el que se inscribe su utilización, los sistemas categoriales deben ser cuanto más generalizables mejor, ya que, al fin y al cabo, se ha invertido mucho tiempo y esfuerzo en su construcción.

Aunque la recomendación anterior debe ser relativizada en función del tipo de registro categorial concreto de que se trate. Así, de entre los tres propuestos en esta obra (sistemas de categorías, listas de control y escalas de estimación), cabe establecer una diferencia de grados entre los sistemas de categorías y los otros dos. Si bien los tres comparten muchos aspectos en común, los sistemas de categorías son mucho más exigentes en cuanto a su elaboración, su validación y su forma de registrar la conducta (generalmente, estableciendo intervalos breves a fin de aproximarse más exactamente a la «unidad» de conducta). Puede decirse de ellos que son verdaderos intentos de medir la conducta por medio de la observación. Por lo tanto, entendemos que en ellos la generalizabilidad (o capacidad de ser aplicados a muchas situaciones y contextos) es fundamental en el sistema. Sin embargo, una lista de control o una escala de estimación, que no son tan rigurosas y exigentes en los aspectos antes mencionados, pueden realizarse para fines concretos y puntuales, por lo que su aplicabilidad puede ser más restringida.

3.1.2. Proceso de elaboración de sistemas categoriales

Ya se indicó en el apartado anterior que la elaboración de un sistema categorial pasa por la elaboración previa de un marco teórico o estructura conceptual que sugiere las conductas a observar y las define operacionalmente. Así, si a través de nuestro sistema pretendemos representar la realidad, debemos garantizar una correspondencia entre ésta y nuestras categorías, aportando definiciones operativas de las mismas.

Para ello es necesario un conocimiento previo del tema que deseamos someter a observación, conocimiento que va a generarse tanto a partir de la lectura de otros

trabajos de investigación y diagnóstico, como de observaciones exploratorias previas que sugieren posibles conductas relevantes a incluir en el sistema.

A continuación vamos a sintetizar las tareas asociadas a la construcción de sistemas categoriales, poniendo un ejemplo de cada una de ellas:

1) Finalidad de la observación. Implica plantearse cuál es la pregunta a la que se quiere responder a través de la observación o cuál es el problema a resolver.

EJEMPLO: Nos interesa conocer si un alumno atiende en las situaciones de enseñanza-aprendizaje que se dan dentro del aula.

2) Marco teórico-conceptual de referencia. ¿De qué supuestos partimos en el estudio de ese fenómeno?, ¿cuáles son las creencias o teorías de las que se parte?, ¿cuáles son las condiciones en las que ese fenómeno se manifiesta?, ¿bajo qué formas se manifiesta?, ¿cómo?, etc.

EJEMPLO: Tendremos que definir en qué consiste la atención, cuáles son los procesos implicados en ella, cuáles son los factores y condiciones que benefician y perjudican los comportamientos atencionales, etc. Tendríamos que indicar también que la atención del alumno puede darse ante explicaciones del docente, o también ante los comentarios de compañeros, o en la realización de actividades individuales y grupales, etc. Ante esta última cuestión, tendríamos que decidir si nos interesan todas estas situaciones en las que se puede manifestar la atención o sólo algunas de ellas (por ejemplo, podemos centrarnos sólo en la atención del alumno ante las explicaciones del profesor).

3) Definición del objeto de observación. Se trata ahora de definir con claridad qué se quiere observar. Esto, que puede parecer una actividad sencilla y obvia, no lo es tanto, ya que hay que definir con bastante precisión y sin ambigüedades el fenómeno que se desea observar. En palabras de KERLINGER (1975), hay que proporcionar una definición operacional de la conducta o conductas a observar, lo que implica definirla/s en términos conductuales.

EJEMPLO: En relación con la atención en el aula, tendríamos que explicitar y definir un conjunto de conductas observables que sean verdaderamente indicativas de que un alumno está atendiendo. Una conducta que podría ser indicativa de atención es que el alumno mire al profesor mientras habla. Ahora bien, puede estar mirando al profesor y estar pensando en otra cosa. Por lo que la conducta definida en esos términos («mirar al profesor»), no es necesariamente indicativa de observación. Quizá «seguir visualmente al profesor con la mirada si éste se mueve o indica algo» sea más precisa a la hora de caracterizar un comportamiento que denota atención.

4) Definición de las categorías incluidas dentro del sistema. Estas categorías derivan de la definición operacional del objeto de observación. Ahora hay que decidir qué categorías se van a incluir en el sistema y acotar, en cada una de ellas, su significado y

su campo de aplicación. Las categorías finalmente formuladas deben ser (ANGUERA, 1991; BAKEMAN y GOTTMAN, 1989; GARCÍA JIMÉNEZ, 1997):

— *Objetivas*: referirse únicamente a características observables de la conducta, evitando referencias a intenciones, estados internos y otros acontecimientos privados.

— *Claras*: no deben ser ambiguas y deben ser fácilmente comprendidas y memorizadas, de tal forma que no haya dudas al atribuir los casos a las categorías.

— *Completas*: deben incluir los límites de la conducta de forma que el observador discrimine entre ésta y otras conductas relacionadas.

— *Mutuamente excluyentes*: de forma que la asignación de un caso a una categoría impida que pueda ser asignado a otra diferente. Las categorías no deben solaparse entre sí, por lo que a cada comportamiento se le asignará una y sólo una categoría.

— *Exhaustivas*: en tanto que las categorías que describen un fenómeno deben agotar todas las posibilidades del mismo. Así, a cualquier comportamiento del ámbito objeto de estudio debe poder asignársele una de las categorías. De esta manera, no deben quedar casos o manifestaciones del fenómeno que no puedan ser asignados a una de las categorías existentes (aunque pueden existir sistemas no exhaustivos).

— *Homogéneas:* esto es, las categorías deben guardar una relación lógica tanto con la variable categorizada como con cada una de las demás. Esto es obvio si tenemos en cuenta que son operativizaciones de un mismo fenómeno.

Para facilitar que se reúnan estos requisitos, debemos incluir algunas indicaciones cuando formulamos las categorías:

— Un nombre.

— Una definición general de su contenido.

— Una presentación de casos críticos e indicaciones de cómo proceder ante ellos.

— Un ejemplo, o más, de conductas integrables en la categoría.

Asimismo, las categorías deben definirse de forma que se contemplen todos sus matices y deben estar acompañadas de ejemplos y contraejemplos para que su especificación sea mayor (ANGUERA, 1991; LOSADA, 1999).

Dado que aquí nos hemos decantado por sistemas con categorías moleculares que soportan un bajo nivel de inferencia, conviene recordar que, generalmente, la presencia de adjetivos y adverbios en las definiciones suele conllevar un mayor nivel de inferencia (por ejemplo, «hacer la tarea bien» puede llevar a que cada observador tenga una concepción previa de qué es hacer una tarea «bien» y «mal»). BLANCO y ANGUERA (1991) plantean además que todas las categorías de un sistema han de situarse en el mismo nivel de descripción, de tal forma que no debe haber categorías que se sitúen en distintos niveles del continuo molaridad-molecularidad.

En definitiva, esta fase es realmente compleja porque implica dar una definición exahustiva de las conductas a observar, especificando los límites de cada una de ellas y los matices de significado que las diferencian. No se trata de una tarea a realizar de forma rápida, sino que exige una reflexión seria por parte de quien elabora el sistema, a fin de garantizar que, una vez diseñado, no va a mostrar dificultades cuando se aplique en situaciones reales. Así pues, lo aconsejable es tomarse todo el tiempo que sea necesario para su construcción.

EJEMPLO: Tomando como referencia la conducta antes enunciada («seguir visualmente al profesor con la mirada si éste se mueve o indica algo»), el proceso de elaboración de la categoría podría hacerse así:

Categoría: «sigue con la mirada al profesor»

Esta categoría incluye todos los comportamientos visuales del alumno que denotan que presta atención a las explicaciones del profesor. Conductas de este tipo que estarían incluidas en la categoría serían: volver la vista si el profesor señala algo en la pizarra, seguir visualmente los movimientos y desplazamientos del docente, volver la cabeza cuando el profesor ha invitado a un compañero a hablar... Otras conductas como asentir con la cabeza ante las explicaciones del profesor o mostrar expresiones de duda y desconcierto, no se codifican en esta categoría, ya que existe una más adecuada para este tipo de conductas («expresiones faciales y movimientos corporales de reacción ante la explicación»).

5) Selección del número de categorías que compondrán el sistema. Además de la forma en que se describen y definen las categorías, es preciso prestar también atención al número de categorías que componen el sistema. Algunos autores consideran que los sistemas de categorías deben tener entre 8 y 14 conductas, pero esto va a depender de las limitaciones del observador, de la complejidad del fenómeno a observar y de las situaciones en que se observa. En realidad, un sistema con muchas conductas es adecuado en los períodos iniciales, con una finalidad exploratoria, pero no tanto para comprobar hipótesis (como, por ejemplo, cuando deseamos estimar los efectos de un tratamiento o programa educativo). Cuantas más categorías tenga un sistema, más difícil resultará diferenciar los matices entre cada una de ellas y, por tanto, existirá menor acuerdo entre los observadores.

6) Otros aspectos a tener en cuenta en la elaboración del sistema categorial. Una vez que las categorías y su número resultan satisfactorias, es conveniente plantear algunas cuestiones que contribuirán a la utilidad del sistema y a su facilidad de uso. Estas cuestiones se refieren a incluir categorías del tipo:

— *Categoría nula*, que se usa para indicar la ausencia de conductas que se consideran relevantes en el sistema.

— *Categoría ficticia,* para cuando dos conductas del sistema tengan lugar de forma contigua.

— *Categorías residuales,* que, durante el proceso de elaboración del sistema, tendrían el carácter de categorías provisionales, pero que por su frecuencia o relevancia pueden ser incluidas en la reelaboración del sistema.

7) *Preparación de la hoja para el registro*. En este punto se trata ya de materializar el sistema de observación en un formato que facilite su aplicación en situaciones reales. Se trata de poner por escrito las categorías del sistema y de preparar unas tablas de doble entrada en las que se van a ir registrando su aparición, secuencia, frecuencia, intensidad... Si el sistema es muy complejo (especialmente los sistemas de categorías), resulta más cómodo para el uso separar ambas partes, elaborando, por un lado, un resumen con la información relativa a las categorías (que, probablemente, tras unas pocas sesiones acabarán siendo memorizadas por el observador) y, por otro, la tabla en la que se va a realizar la codificación.

Hasta aquí, el proceso descrito se ha referido estrictamente a la elaboración del sistema categorial. Es decir, a definir las conductas que van a ser objeto de observación y preparar una hoja de registro acorde con el sistema diseñado. No obstante, y como se explicitó en el apartado anterior, hay otras cuestiones a abordar por parte del constructor del sistema, cuestiones que se refieren a cómo se debe aplicar el mismo: la duración de la observación, su frecuencia, el establecimiento de intervalos, la selección de situaciones, de sujetos, etc.

Tanto en su presentación a los observadores, como al público en general, un sistema categorial no debe limitarse a enunciar el conjunto de categorías que se proponen para la observación. Debe también especificar en qué condiciones puede ser usado, durante cuánto tiempo debe usarse, cuál es la duración óptima de las sesiones e intervalos de observación, así como cuantas consideraciones sean necesarias para facilitar su adecuada aplicación. Si un grupo de docentes quisiera aplicar un sistema categorial ya elaborado para observar la atención del alumno en el aula, debe saber si dicho sistema se puede utilizar en cualquier situación de aula o sólo en las clases de matemáticas o sólo cuando se producen explicaciones del profesor, etc. Debe, asimismo, tener pautas que le orienten sobre las cuestiones relativas al tiempo. En caso contrario, los docentes podrían usarlo durante demasiado poco tiempo o en intervalos poco adecuados, etc. Por lo tanto, explicitar las condiciones y formas en que el sistema categorial puede y debe usarse es una tarea más dentro del proceso general de elaboración del instrumento.

De cualquier forma, este proceso no debe darse nunca por concluido o acabado, ni tampoco el sistema categorial diseñado debe entenderse como definitivo. Muy al contrario, el proceso de construcción de un sistema de categorías es lento y se va optimizando progresivamente hasta que se adapta adecuadamente a la situación para la cual fue elaborado. Esta afirmación no sólo es válida para el más sofisticado sistema de categorías, sino también para la lista de control o escala de estimación que se diseña para procesos diagnósticos concretos.

Cuando se depura un sistema de categorías, un aspecto que nos da mucha información de cara a su mejora es la frecuencia con que tienen lugar y se registran las conductas. Si encontramos que para algunas categorías esta frecuencia es mínima, quizá sea conveniente replantear la utilidad de la misma o reflexionar sobre el marco teórico que la ha sugerido. Al mismo tiempo, una frecuencia demasiado alta puede indicar que la categoría en cuestión no discrimina convenientemente, por lo que es necesario revisarla en los mismos términos.

Resulta muy útil preguntar a los observadores si han tenido dificultad para discernir entre dos o más categorías, haciendo que nos pongan ejemplos de conductas concretas que se sitúan en una zona borrosa. ANGUERA (1991) habla de la *borrosidad categorial*, en tanto que la membrecía o pertenencia de la conducta a una categoría a veces es una cuestión de grado y, por tanto, resulta difícil tomar la decisión sobre cómo registrarla. La aportación al respecto de la persona que está usando el sistema es de gran ayuda para identificar estos casos.

En general, los usuarios de un sistema deben estar atentos a las inadecuaciones de éste, a fin de conseguir una mayor adaptabilidad del instrumento. Esta misma idea es expresada sabiamente por BAKEMAN y GOTTMAN (1989:61), quienes afirman que «un sistema de categorías puede evolucionar en la medida en que es utilizado por codificadores inteligentes».

3.1.3. Tipos de sistemas categoriales

Los sistemas categoriales se caracterizan por centrarse en unidades de observación especificadas *a priori.* Anteriormente hicimos referencia a tres tipos de sistemas categoriales: los sistemas de categoría, las listas de control y las escalas de estimación. En el anexo 3 presentamos un ejemplo de cada una de estas técnicas de registro que ahora pasamos a comentar.

Los **sistemas de categorías** dan su nombre al conjunto de sistemas categoriales para la observación. Se caracterizan por constituir verdaderos modelos en los que se concreta la explicación de un fenómeno o conducta. En ellos, la observación está basada en una fundamentación conceptual sólida y existe una mayor preocupación por garantizar la fiabilidad y validez del sistema finalmente diseñado. Asimismo, el establecimiento de intervalos de tiempo es una cuestión esencial en ellos. En función de estos intervalos, será posible obtener muestras representativas de la conducta y una estimación más certera de su frecuencia o duración.

Todos los aspectos a considerar en los sistemas categoriales (molaridad-molecularidad, aplicabilidad a distintas situaciones, etc.) son abordados exhaustivamente en su proceso de elaboración, el cual se ajusta más rigurosamente a la propuesta de pasos que hemos realizado con anterioridad. Como consecuencia de ello, están próximos a ser auténticos procedimientos de medición de la conducta con base en la observación. Si tuviéramos que hacer una analogía, los sistemas de categorías, en el campo de la

observación, son equivalentes a los tests, mientras que las listas de control y las escalas de estimación se asemejarían más a un cuestionario. A través de estos últimos, se puede cuantificar la conducta, pero no necesariamente ofrecer una medida válida y fiable a través de ellos.

Las **listas de control** también contienen un conjunto de conductas establecidas *a priori* cuya ausencia o presencia se registra durante el período de observación. En este tipo de registro, si bien es necesaria una fundamentación conceptual que dé sentido a las categorías seleccionadas para la observación, el proceso de elaboración no es tan riguroso como en el sistema de categorías que, como dijimos, intenta aproximarse a la medición de la conducta.

Asimismo, en ellos la observación suele basarse en el acontecimiento (EVERTSON y GREEN, 1989), es decir, se trata de registrar si aparece o no la conducta durante el período de observación prefijado, sin necesidad de establecer intervalos de tiempo. Son útiles cuando queremos observar conductas poco frecuentes o muy diversificadas en sus manifestaciones. El observador se limita a indicar si tales conductas se dan o no en el sujeto o grupo observado, sin indicar su magnitud, frecuencia o duración. Cuando una determinada conducta aparece con alta frecuencia, no es adecuada su inclusión en la lista —en cuyo caso, sería recomendable una escala de estimación—, ya que las listas de control están pensadas para conductas críticas o relevantes cuya aparición o no, en un determinado contexto, nos interesa comprobar.

En opinión de algunos autores (BLANCO y ANGUERA, 1991), las listas de control sólo exigen una pre-planificación de las conductas por separado y una ordenación lógica de las mismas. Sin embargo, nosotros pensamos que en su construcción, son válidas todas las recomendaciones sugeridas en el apartado anterior, ya que como sistemas categoriales que son, se rigen por los mismos principios y proceso de elaboración de este tipo de registros. No obstante, cuando las categorías a observar sean conductas que forman parte de un mismo proceso, conviene ordenarlas según la secuencia en que dicho proceso suele operar (CABRERA y ESPÍN, 1986). Por ejemplo, si se trata del comportamiento a desarrollar por un trabajador en una jornada laboral, resultará más fácil hacer la observación si la lista de control tiene las categorías más o menos ordenadas según la secuencia habitual de trabajo.

Las escalas de estimación —también llamadas escalas de apreciación y *rating scales* — se utilizan para cuantificar las impresiones que se obtienen a partir de la observación y exigen juicios cuantitativos sobre el grado o intensidad de la conducta. En esencia, consisten en un conjunto de características o comportamientos a observar y valorar de acuerdo con algún tipo de jerarquía (CABRERA y ESPÍN, 1986). Los puntos a lo largo de cada escala representan distintos grados del atributo que está bajo observación. Generalmente, en ellas, el registro se produce después de la observación, por lo que presentan un alto grado de inferencia (BLANCO y ANGUERA, 1991).

A diferencia de las listas de control, las escalas de estimación son adecuadas para conductas con alta frecuencia de aparición o poco diversificadas en sus manifestaciones. Así, ante conductas frecuentes, no cabe plantearse si aparecen o no, sino más bien en qué grado, o con qué intensidad se manifiestan.

Existen diversos tipos de escalas de estimación (BLANCO y ANGUERA, 1991; KERLINGER, 1975; PÉREZ JUSTE, 1991), entre los cuales resaltamos:

— *Escalas descriptivas o verbales*. En ellas, los rangos de la escala se expresan mediante expresiones verbales más o menos descriptivas. En el cuadro 2 aparece un ejemplo de este tipo de escalas.

— *Escalas numéricas.* En las que se asignan números al grado o intensidad de la conducta, de acuerdo con una equivalencia previamente establecida. En este caso, es preciso indicar si la escala es creciente (el valor 1 es la mínima intensidad) o decreciente (el valor 1 es la máxima intensidad).

— *Escalas gráficas*. Donde a través de líneas o barras se sitúan las frases descriptivas o los valores numéricos, permitiendo una mayor precisión en la selección del rango.

Como sistema categorial, en su construcción se siguen los mismos pasos expuestos en el apartado anterior, aunque no necesariamente de forma tan rigurosa y exhaustiva. Además de las decisiones respecto a las categorías a incluir en la escala, es conveniente tener en cuenta ciertos aspectos en la elección y formulación de los rangos, especialmente cuando se trata de escalas de estimación descriptivas. Entre los aspectos propuestos por PÉREZ JUSTE (1991), resaltamos los siguientes:

— Elección de la dimensión de la conducta que se va a considerar: intensidad, duración, permanencia, arraigo, etc.

— Decisión sobre el número de rangos que tendrá la escala.

— En el caso de las escalas descriptivas, además hay que procurar que los rangos sean exhaustivos, es decir, que no queden fuera de la escala posibles «grados» relevantes para valorar la conducta; y exclusivos, en tanto no exista superposición entre ellos. Además es preciso que exista unidad de enfoque en los rangos, estando todos ellos en una misma dimensión (duración, arraigo, intensidad...): por ejemplo, sería incorrecta una escala con los rangos MUCHO, BASTANTE, A VECES, NUNCA; en este caso, MUCHO y BASTANTE indican intensidad de la conducta o arraigo, mientras que A VECES y NUNCA indican frecuencia. Se están pidiendo, por tanto, dos tipos de valoraciones distintas no necesariamente comparables.

CUADRO 2

EJEMPLOS DE DISTINTOS TIPOS DE ESCALA DE ESTIMACIÓN

Ejemplo de escala de estimación descriptiva:
CATEGORÍA: Realiza las actividades de casa solicitadas por el profesor.
RANGOS: Nunca Pocas veces A veces Casi siempre Siempre

Ejemplo de escala de estimación descriptiva con información adicional:
CATEGORÍA: Realiza las actividades de casa solicitadas por el profesor.
RANGOS: Siempre, sólo en muy raras ocasiones no trae hecha la tarea de casa.
A veces, sólo las hace si la tarea le gusta o le resulta fácil.
Nunca, en contadas ocasiones ha traído tareas realizadas en casa.

Ejemplo de escala de estimación numérica:
CATEGORÍA: Realiza las actividades de casa solicitadas por el profesor.
RANGOS: 1 2 3 4 5 6

Ejemplo de escala de estimación gráfica:
CATEGORÍA: Realiza las actividades de casa solicitadas por el profesor.
RANGOS: 1 2 3 4 5 6 7

Dado que en la escala de estimación está implícita la valoración de la conducta (emisión de un juicio de valor sobre su intensidad, duración...), en ocasiones, se considera una técnica subjetiva. Si bien puede tener múltiples aplicaciones en el campo educativo, conviene conocer algunos de los factores que afectan a la calidad de la valoración que en ella se emite. Así, los principales inconvenientes que presenta derivan de los *errores y efectos* que son habituales en su aplicación:

— *Error de tendencia central* o tendencia a evitar los extremos de la escala. Suele manifestarse cuando los evaluadores desconocen los objetos que están clasificando, usando la opción central cuando no tienen un juicio muy claro.

— *Error de lenidad*, o de indulgencia. Es la tendencia a conceder puntuaciones altas a todos los sujetos.

— *Error de severidad* o tendencia a puntuar bajo.

— *Efecto de halo*. Este efecto, descubierto por Thorndike, hace referencia a que el juicio emitido por el observador deriva de la impresión general que tiene de lo que está observando (KERLINGER, 1975). También se manifiesta en la tendencia a encadenar unas categorías con otras al influir la estimación que se da a un elemento en la que se da al siguiente. Un ejemplo de este efecto es el que se da cuando un profesor valora las conductas de un alumno negativamente por el hecho de que es

revoltoso. O cuando un evaluador emite valoraciones positivas de un centro por el mero hecho de que lo han atendido bien.

Pese a estos inconvenientes, tanto las escalas de estimación como las listas de control son relativamente fáciles de construir y también muy fáciles de aplicar. En su favor, hemos de decir también que ambas tienen multitud de aplicaciones en el campo educativo, tanto como instrumento para la evaluación del alumnado, como para la evaluación de otras dimensiones. Ambos instrumentos se pueden usar, por ejemplo, para valorar contenidos actitudinales y procedimentales, más difíciles de estimar a través de pruebas escritas. Pueden aplicarse también a la valoración de las producciones y tareas del alumno. En este caso, se basan en la observación, ya no de un sujeto, sino de productos (trabajos monográficos, cuadernos de actividades, composiciones ...) realizados por el mismo.

Asimismo, constituyen una estrategia muy adecuada para evaluar otros objetos o entidades. Se suelen utilizar, por ejemplo, en la evaluación del diseño de programas. En estos casos, se recogen en ellas una serie de indicadores de la calidad del diseño y se someten al juicio de personas expertas en la temática que aborda el programa. Pueden aplicarse también a la evaluación de instituciones siguiendo un planteamiento parecido. Por lo tanto, su utilidad en el campo educativo está fuera de toda duda.

3.2. Observación no sistematizada

En apartados anteriores hemos caracterizado esta modalidad de observación como una forma no controlada de observar, en la que el observador sólo parte de ideas muy generales sobre lo que quiere observar, sin predefinir las conductas o categorías en las que va a centrar su atención. Al no existir puntos o ejes en torno a los cuales se especifica la observación a realizar, existen múltiples formas de hacer observación no sistematizada, y múltiples tipos de registros que acompañan a este tipo de observación. Al mismo tiempo, entre estos registros no resulta fácil delimitar un patrón común, como sucedía en los sistemas categoriales, por lo que, en este apartado, daremos algunas ideas generales sobre esta modalidad de observación y pasaremos después a centrarnos en la especificidad de cada registro.

EVERTSON y GREEN (1989) distinguen entre dos formas de observación no sistematizada o abierta: los *sistemas descriptivos*, que buscan descripciones abundantes y detalladas de los fenómenos o conductas, y los *sistemas narrativos*, en los que se adopta una perspectiva cronológica que pretende identificar patrones de comportamientos. En este último tipo de registros, la implicación del observador es una de las características más señaladas; sus impresiones sobre lo que observa serán de gran importancia en el proceso de observación. Aquí no vamos a utilizar esta diferenciación, ya que pensamos que ambos sistemas de observación comparten aspectos comunes, al tiempo que existe gran diversidad en las formas concretas en que se lleva a la práctica la observación no

sistematizada, por lo que consideramos más conveniente centrarnos en cada tipo de registro concreto.

Sin embargo, la aportación de estas autoras nos permite llamar la atención sobre la vinculación existente entre las diversas modalidades de observación que planteábamos al inicio de este capítulo. Nuevamente queremos recordar que tomar como criterio el grado de sistematización de la observación para presentar los diferentes registros observacionales, no implica que toda observación no se deba caracterizar además por el nivel de participación del observador, y por el hecho de que éste se observe a sí mismo (auto-observación) o a otras personas, situaciones o cosas (hetero-observación). Así, en muchos de los registros que presentaremos con posterioridad, se verán claramente reflejadas las tres modalidades de observación.

La observación no sistematizada o abierta, al no estar enfocada en unos eventos o comportamientos prefijados, proporciona descripciones abundantes y copiosas de las conductas en toda su complejidad y extensión, encuadrándolas en el contexto y situación en que surgen y se manifiestan. Según IBÁÑEZ (1993), resulta especialmente útil cuando se desea una primera impresión de la conducta del sujeto en uno o varios *settings*, como precursor de un sistema de categorías y como medio para generar hipótesis sobre las variables que controlan la conducta. Nosotros pensamos que también es adecuada cuando se desea comprender una realidad, más que explicarla o cuantificarla.

Entre las ventajas de los registros no sistematizados, IBÁÑEZ (1993) sugiere las siguientes: proporcionan una descripción de la conducta y su contexto, mantienen íntegra la secuencia de eventos y suministran información valiosa sobre conductas críticas y sus circunstancias.

Sin embargo, resulta difícil establecer criterios para un tipo de observación que se define como no sistematizada. Aunque dependerá en gran medida del problema que suscita el diagnóstico, ANGUERA (1988) indica algunos puntos generales que permiten optimizar este tipo de observaciones, entre los cuales hemos seleccionado los siguientes:

— La observación debe intentar establecer el complejo de relaciones del comportamiento observado, la situación y el sistema socio-cultural.

— La secuencia observada es desatada por factores propios de la situación, por lo que es necesario determinar con la mayor precisión posible las condiciones iniciales bajo las que suelen aparecer las conductas.

— Es necesario también determinar la estructura de posiciones de la situación, ya que en muchas ocasiones el comportamiento que observamos es el de personas concretas que desempeñan roles concretos.

Para comprender una conducta, un fenómeno o una secuencia de los mismos, es necesario conocer el mayor número de detalles posibles sobre los aspectos indicados más arriba, con el fin de que sea posible inferir la concurrencia de una determinada

constelación de factores a partir de características de comportamiento directamente observables.

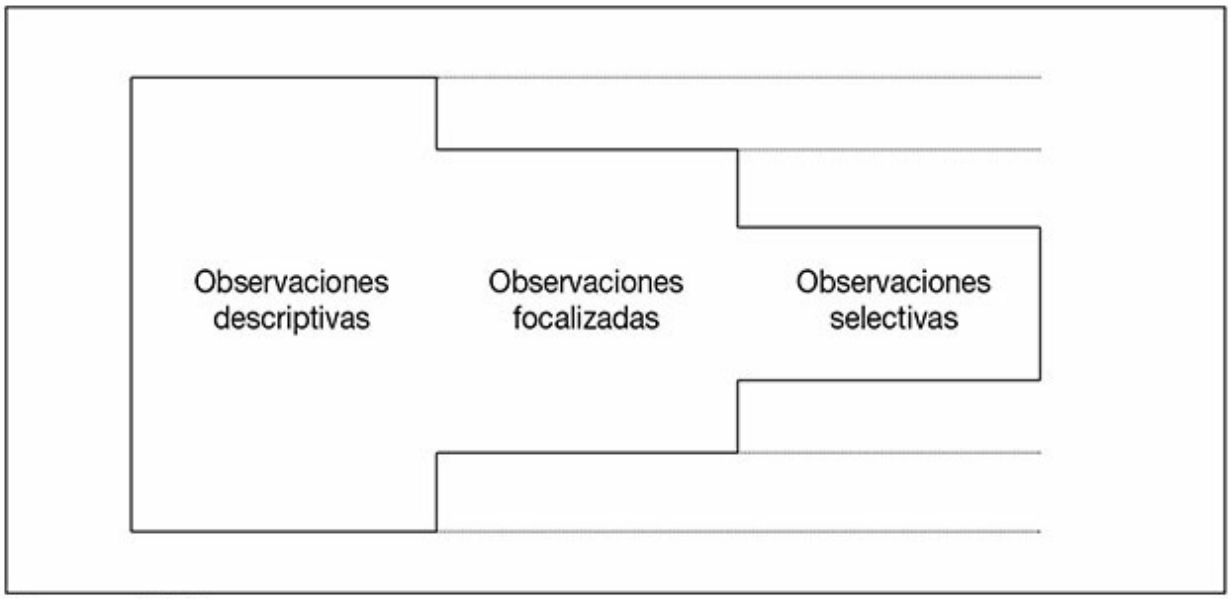
Dadas todas estas características, en la observación no sistematizada suele ser habitual que la selección de muestras se acoja a los límites naturales dentro de los que se manifiestan las conductas, acontecimientos y procesos de interés para el diagnóstico. Así, tanto la duración total de la observación, como la duración y distribución de los períodos de observación se ven condicionadas por la cuestión a la que se pretende dar respuesta y el marco natural en que tienen lugar los eventos estudiados. Si tuviéramos interés en conocer cómo participan los estudiantes en la dinámica universitaria, tendríamos que observar en las distintas situaciones en las que esta participación se puede dar (dentro de una clase, en una sesión de Claustro, en una Junta de Facultad...); en todos los casos, la observación se realizaría durante la duración del evento (clase, sesión del Claustro, Junta de Facultad, etc.), sin segmentar artificialmente el marco temporal. Asimismo, la duración de la observación debería ser tal que permitiera captar todas las facetas de la participación universitaria. Un mes, por ejemplo, sería una cantidad de tiempo insuficiente. En un plazo tan corto puede que no tengan lugar determinados acontecimientos que serían interesantes para nuestro estudio (por ejemplo, no es lo mismo la participación en épocas de exámenes que al principio de curso; los distintos actos académicos y culturales no se realizan según una secuencia rutinaria —una vez a la semana, al mes...—). Por lo tanto, el marco natural en que tienen lugar los eventos de interés será la referencia más importante a la hora de decidir sobre la selección de muestras en este tipo de observación.

Si en los sistemas categoriales las unidades de observación eran las categorías, en los sistemas abiertos, las unidades de observación abarcan múltiples aspectos de la conducta, ya que lo que se pretende es reflejar ésta en toda su complejidad y extensión. El registro de la observación tiene que dejar constancia de la ocurrencia natural de las acciones, acontecimientos, conductas... sin tratar de filtrar lo que ocurre, sino recogiendo los acontecimientos tal cual ocurren y sin separarlos del contexto-situación en que tienen lugar.

El proceso de observación, no obstante, suele seguir en ocasiones la denominada *estrategia del embudo* (SPRADLEY, 1980) (ver fig. 2). Así, al comienzo de la observación, el observador no tiene muy claro qué debe observar, por lo que no se centra en cuestiones específicas, sino que observa de forma general todo lo que ocurre (*observación descriptiva*), siendo muy amplias las unidades de observación. Conforme avanza el proceso, determinadas conductas comienzan a mostrarse como más relevantes en relación con el problema que suscita la observación, por lo que el observador tiende a centrarse más en ellas (*observación focalizada*). Finalmente, el proceso puede culminar con *observaciones selectivas*, en las que el observador se centra en observar aquello que le permite contrastar sus hipótesis.

FIGURA 2

ESTRATEGIA DEL EMBUDO:
ETAPAS EN LA OBSERVACIÓN NO SISTEMATIZADA



(SPRADLEY, 1980)

Si bien uno de los principales peligros de este tipo de observación es la arbitrariedad en las observaciones de distintos observadores sobre el mismo objeto (ANGUERA, 1988), es de destacar que el observador es el instrumento primordial de la observación no sistematizada. Así, el registro de la información depende generalmente del sistema perceptivo de éste y de su habilidad para capturar en lenguaje ordinario lo que observa.

Otra crítica recibida por este tipo de observación —aunque, en realidad, se trata de una crítica general a todos los sistemas observacionales— se refiere a que las personas modifican su comportamiento habitual por el hecho de sentirse observadas. En una observación no sistematizada y participante, esto puede originar que la presencia del observador altere más o menos significativamente el contexto y personas a las que estudia. No obstante, nosotros pensamos que tras unos primeros momentos en que esta reactividad a la observación puede tener lugar, las personas se acostumbran a ser observadas y comienzan a comportarse de forma normal. Esto mismo defienden GARCÍA FERRANDO y SANMARTÍN (1986:114) cuando plantean que:

> «los actores no pueden dejar de vivir sus vidas y estar siempre, durante tan largo plazo, alterando su conducta. Terminan acostumbrándose a nuestra presencia y progresivamente decaen tanto su curiosidad por el observador, como su desconfianza».

En el campo educativo, la observación no sistematizada tiene múltiples aplicaciones. Permite acceder a las realidades educativas sin filtrar previamente la información que de ellas se va a extraer. De esta forma, se pueden conocer y comprender los factores contextuales que inciden en una determinada situación. Más allá de ofrecer descripciones

superficiales de las conductas, permite profundizar en los factores situacionales y escenarios en que éstas tienen lugar, así como en los comportamientos peculiares e idiosincráticos de cada sujeto, lo que será de gran valor a la hora de diseñar propuestas formativas adaptadas a cada sujeto. Según HERNÁNDEZ DE LA TORRE (1997), la observación aplicada al entorno educativo debe tener un marcado carácter naturalista.

3.2.1. Tipos de registros de observación no sistematizada

En general son muchos los posibles tipos de registros que se pueden utilizar en la observación no sistematizada. Sea cual sea la forma de registro elegida, es conveniente en todos los casos seguir algunos consejos como: identificar tanto *quién* observa como *a quién* se observa; quién es el emisor de la conducta; cuáles son las conductas observadas y qué factores situacionales parecen afectarle; cuáles son los escenarios en que tiene lugar, cómo afectan éstos a la conducta, etc.

Durante la situación de observación, es necesario llevar algún tipo de registro de lo que se observa para evitar la pérdida de información y la distorsión de la memoria. Incluso en los casos en que resulta muy difícil —o contraproducente por las reticencias o la reactividad que puede generar—, se debe tomar algún tipo de notas en las que, al menos, figuren expresiones clave o palabras significativas que faciliten el recuerdo posterior de los acontecimientos. Con posterioridad a la situación de observación, estas notas deberán ampliarse tan pronto como sea posible, con todos los detalles que se recuerden de la observación. Las notas tomadas durante la observación ayudarán a reconstruir el fenómeno o conducta en su globalidad (GARCÍA FERRANDO y SANMARTÍN, 1986).

De acuerdo con lo anterior, podemos clasificar los registros de observaciones no sistematizadas en dos grandes grupos: por un lado, aquellos en los que los acontecimientos y conductas se recogen en vivo, durante la observación (*registros in situ*) y, por otro, aquellos en que el registro se produce una vez finalizada la sesión de observación (*registros a posteriori*). A continuación, presentamos los principales tipos de registros en cada uno de los grupos mencionados.

3.2.2. Registros *in situ*

Entre los diferentes tipos de registros en vivo que se pueden realizar durante la observación no sistematizada, en este apartado nos ocuparemos de las notas de campo, los registros anecdóticos y los registros de muestras.

Las **notas de campo** son registros efectuados en lenguaje cotidiano acerca de lo que se observa. Consisten en apuntes o notas breves que se toman durante la observación para facilitar el recuerdo de lo observado. ANDER-EGG (1982) se refiere a ellas como la libreta que el observador lleva siempre consigo, a fin de registrar sobre el terreno todas las informaciones, fuentes de información, expresiones, hechos, croquis, etc. que pueden ser de interés en su trabajo. Pueden considerarse como un tipo de registro casi esencial

en la observación no sistematizada, ya que, independientemente de que se usen otros registros *in situ,* casi siempre las observaciones de este tipo suelen registrarse en notas de campo.

En la libreta de notas, por otra parte, se recogen muchas informaciones que serán de gran importancia para que el observador vaya centrando y focalizando su observación. En ella es posible anotar todos los interrogantes o cuestiones que surgen durante la observación, a fin de tenerlos en cuenta en futuras sesiones. Asimismo, como plantean EVERTSON y GREEN (1989), en ellas se recoge también información sobre cuestiones metodológicas (referidas a las decisiones que toma el observador), personales (referidas a las reacciones personales del observador) y teóricas (referidas a los vínculos con la teoría y con los patrones observados).

Por lo general, al tomarse mientras se observa, resulta difícil hacer un registro exhaustivo de todas estas informaciones, por lo que resulta habitual que en esta libreta se recojan ideas, palabras o expresiones clave, dibujos, esquemas, etc. Todas estas anotaciones facilitarán el recuerdo posterior y permitirán la construcción de unas notas ampliadas o de un diario de observación en los que se hará un registro más detallado de los acontecimientos *a posteriori.*

Los **registros anecdóticos**, como su nombre indica a la perfección, son registros de anécdotas, o breves descripciones de una conducta o evento. Estas anécdotas suelen hacer referencia a comportamientos poco usuales. Así, sería absurdo hacer registros anecdóticos de todos los comportamientos que se dan en una sesión de observación, para ello hay otros procedimientos más eficaces como las notas de campo. Por tanto, a través de los registros anecdóticos se recogen acontecimientos que adquieren una especial significatividad para quien realiza la observación. Normalmente, estos acontecimientos o conductas serán significativos o relevantes en relación al fenómeno objeto de estudio o problema que suscita la observación.

Al centrarse en situaciones significativas, algunos autores prefieren denominarlos como *registros de incidentes críticos*, ya que los eventos a partir de los cuales se producen los registros deben ser críticos, es decir, que tengan la suficiente complejidad y significación para que a partir de ellos se pueda inferir algo sobre la persona o grupo objeto de observación.

En cualquier caso, es importante que el contenido de la anécdota se relacione, según PÉREZ JUSTE (1989), con las conductas tanto positivas como negativas. En ocasiones, al aplicar este tipo de registro el observador tiende a seleccionar lo negativo en mayor medida que lo positivo. El valor potencial de la ficha anecdótica reside en su riqueza, su matización en función del contexto y su valor secuencial.

Aunque el formato del registro puede variar según las necesidades de la observación en la que se usa, es adecuado incluir siempre unos datos mínimos como la fecha, hora de observación, observador, descripción del contexto en que tiene lugar el evento, descripción del incidente y valoración del mismo (ver un ejemplo en el cuadro 3). La

mayoría de los autores defienden que es necesario separar la descripción del incidente de la valoración que el observador emite del mismo. PÉREZ JUSTE (1990) y BLANCO (1994) consideran además que es conveniente que la valoración del incidente vaya en el reverso del registro, a fin de separar físicamente la descripción y la valoración de la conducta.

CUADRO 3

EJEMPLO DE REGISTRO ANECDÓTICO

REGISTRO ANECDÓTICO
Fecha: 24 Junio 2001 **Hora**: 12,30h.
Protagonista/s del incidente: Felipe y todo el equipo técnico.
Observador/a: Marisa.
Contexto: Reunión de equipo técnico para la elaboración del plan de trabajo para el curso próximo.
Descripción del incidente: Luisa detecta que el lenguaje utilizado en el proyecto es sexista y sólo utiliza el masculino. Esto provoca la oposición de Felipe que defiende que el masculino puede usarse de forma genérica. Se genera una discusión acalorada sobre el tema y al final se decide someterlo a votación.
Valoración: Siguen existiendo muchas resistencias a abordar el sexismo existente en nuestra Asociación, a pesar de que es uno de los objetivos recogidos en sus Estatutos. Estas resistencias provienen siempre del mismo sector: los técnicos varones con más antigüedad. La discusión de hoy ha sido muy importante porque muestra cómo cada vez hay más personas sensibilizadas con el tema y, por tanto, hay que llegar a acuerdos mínimos.

Aunque aquí hemos incluido el registro anecdótico dentro del grupo de los registros *in situ*, existen opiniones diferentes sobre el hecho de si deben cumplimentarse en el acto o con posterioridad. BLANCO y ANGUERA (1991) y EVERTSON y GREEN (1989), por ejemplo, plantean que se pueden realizar tanto en vivo, como de forma retrospectiva. Otros autores, como BLANCO (1994), recomiendan usarlos *a posteriori,* a fin de que no influya el estado de ánimo del observador en el momento inmediato en que tiene lugar el evento. Nosotros pensamos que, dado que este tipo de registro no reviste excesiva complejidad, se pueden realizar en vivo, durante el proceso de observación; al menos, en lo referido a la descripción del contexto y del incidente, si bien la valoración del evento o conducta puede registrarse *a posteriori*.

Entre las ventajas de este tipo de registro, cabe plantear la facilidad de su cumplimentación. Al ser relativamente poca la información a consignar, y dado el hecho de que sólo se utiliza ante incidentes críticos, son muy fáciles de usar dentro de la observación participante. Sin embargo, hay que tener en cuenta que constituyen un

sistema restringido en el que se registra sólo un segmento específico de la realidad, por lo que deben usarse de forma complementaria a otras estrategias que recojan la observación de forma más secuencial y continuada. Por otra parte, en la medida en que se acumulen muchos registros, la técnica ofrecerá más información, ya que para inferir algo sobre los patrones comportamentales de un sujeto, será necesario tener un repertorio amplio de observaciones sobre el mismo.

Los **registros de muestras** son otro tipo de registros en vivo, más detallados que los registros de incidentes críticos. La finalidad de este tipo de registro es obtener descripciones de acontecimientos que aparecen en forma sistemática e intensiva durante la observación. Se pretende obtener un flujo ininterrumpido de conducta con el mayor detalle que sea posible sin seleccionar ni filtrar lo que ocurre: el observador registra lo que dice y hace un sujeto seleccionado (o un grupo de sujetos), así como la información relativa al contexto. Las conductas de otros individuos sólo se registran en relación con la de la persona que se seleccionó para ser observada. Se procura que el registro sea cronológico, recogiendo los pasos implicados en cualquier acción o conducta dada.

BLANCO y ANGUERA (1991) los denominan *registros de formatos*. Según estos autores, se trata de obtener descripciones de eventos que son registros *on-line,* sistemáticos e intensivos, dado que el observador registra la conducta que ocurre durante un período de tiempo designado de antemano y con tanto detalle como le resulte posible. Estos autores recomiendan además evitar la interpretación de las conductas observadas. Cuando finaliza el registro, el observador segmenta el «flujo de conducta» en episodios que reflejan la acción o situación. Dichos episodios se construyen a través del reconocimiento de conductas que tienen una dirección consistente o que tienen un objetivo definido.

3.2.3. Registros *a posteriori:* el diario

El diario es un tipo de registro observacional que se construye *a posteriori* de la observación. Apoyado en las notas de campo, el observador recoge en él no sólo lo que recuerda, sino también las reflexiones o impresiones provocadas por lo que ha observado. Es un tipo de registro que sirve tanto para la hetero-observación como para la auto-observación (BLANCO y ANGUERA, 1991; LÓPEZ GÓRRIZ, 1997). Así, se puede utilizar para registrar los hechos y conductas de otras personas que se producen en una situación determinada, pero también para narrar las vivencias y sentimientos que esos hechos provocan en el observador. No siempre estas dos dimensiones (auto y hetero-observación) aparecen claramente diferenciadas en el diario, siendo ésta una de sus principales virtudes: «aparecer como un espacio de reflexión escrita en el que se integran y mezclan "narrador" y "cosa narrada", sujeto y hechos, uno mismo y lo que se cuenta». (ZABALZA, MONTERO y ÁLVAREZ, 1986:299).

Entre sus características principales, BLANCO y ANGUERA (1991) señalan que son registros retrospectivos y longitudinales de eventos y sujetos que contienen observaciones repetidas de un individuo o grupo durante un período de tiempo; las

observaciones dependen de las percepciones y el recuerdo del observador; los patrones de conducta se extraen del texto creando entonces unidades inferidas y, finalmente, las unidades pueden identificarse dependiendo de la estructura de la persona que escribe el diario y del propósito que se tenga para conservar estos registros. Asimismo, el marco de referencia de la persona que lo redacta influye sobre lo que se registra y las observaciones dependen de sus recuerdos y percepciones (EVERTSON y GREEN, 1989).

El diario es un recurso bastante costoso en tanto que exige una continuidad en el esfuerzo por describir y narrar los eventos. Para su autor implica tanto una inversión de tiempo —dedicar un tiempo a la escritura tras la observación de los eventos—, como una inversión de esfuerzos —que supone reflexionar sobre lo observado y reconstruirlo verbalmente por escrito—. Ahora bien, una vez adquirido el hábito, las personas que lo utilizan le encuentran gran sentido y utilidad para objetivar y analizar la realidad que está siendo objeto de observación, ampliándose a menudo los objetivos que inicialmente se había propuesto el autor y encontrando nuevos ámbitos de aplicación.

Como registro observacional, el diario presenta la ventaja de captar el carácter longitudinal, cronológico y procesual de los eventos que están siendo observados. A través de él se recoge información de forma continua y sistemática (LÓPEZ GÓRRIZ, 1997). Permite, así, establecer la secuencia de los acontecimientos, dando cuenta de cómo éstos se desarrollan y evolucionan dentro de un marco temporal. A veces, la perspectiva que adopta el autor del diario puede ser retrospectiva (reflexión sobre acontecimientos pasados que se observan con el prisma del momento presente), pero eso no implica que el diario sea una reconstrucción del pasado desde el momento actual (como suele serlo la historia de vida).

Por otra parte, el hecho de ser un registro procesual, que se desarrolla progresivamente en un marco temporal, le confiere la potencialidad de estimular y perfeccionar la capacidad de observación y descripción de las situaciones (LÓPEZ y PADILLA, 1997). Así, conforme el sujeto va observando una realidad —externa o interna a sí mismo— e integrándose en ella, se agudiza su capacidad de observación, es decir, se va haciendo más sensible a multitud de aspectos que en principio le pasan desapercibidos (LÓPEZ y otros, 1996).

BEHAR (1991) plantea que lo espontáneo y subjetivo de las manifestaciones que se encuentran en un diario, junto con su carácter espontáneo confieren un valor inapreciable a los datos procedentes de los mismos. Así, a diferencia de otras estrategias como la autobiografía, las anotaciones realizadas en un diario no surgen todas al mismo tiempo y con carácter retrospectivo, sino que van teniendo lugar día a día, reflejando los cambios que se producen en el mismo momento en que tienen lugar. Sin embargo, cuando el diario es llevado durante un cierto tiempo, el sujeto llega a integrar diversas anotaciones en un todo e incluso a seguir un plan tendente a alcanzar coherencia narrativa.

En el diario, el observador analiza los acontecimientos actuales desde la lógica del momento presente y de acuerdo con su capacidad de análisis de ese momento. Esto no

ocurre en la historia de vida, donde se revisan hechos pasados a la luz desde la óptica del presente. Por eso, «el diario es altamente revelador de la relación entre los sucesos escritos en el mismo y las estructuras cognitivas del autor del diario en que éste es escrito» (HARRÉ y DE WAELE[3], 1979:177, cit. por BEHAR, 1991).

Por otra parte, al ser un registro *a posteriori,* el diario facilita pasar de las emociones que suscitan los eventos observados a la clarificación y racionalización de la situación en lo abstracto, conllevando esto la liberación del peso emocional que despierta lo observado (LÓPEZ GÓRRIZ, 1997).

Entre las limitaciones del diario, BEHAR (1991) señala la tendencia a que se reflejen en él los hechos de especial relevancia, ya sean conflictos o situaciones altamente gratificantes, por lo que rara vez aparecen en él descripciones detalladas de la normalidad. Esto, además de suponer sesgos en la información respecto a diversos aspectos de la vida del sujeto, puede producir saltos temporales en el diario. Por otro lado, debemos tener en cuenta su carácter íntimo y privado, por lo que no siempre resulta fácil acceder a este tipo de información.

Algunas dificultades que aparecen cuando se usa el diario por primera vez son: no saber bien qué registrar y cómo, lo que lleva a dos tendencias distintas, bien a anotar todo lo que ha ocurrido en la sesión de observación, o bien a no anotar muchos eventos relevantes. También suele ser frecuente que, al principio, su autor adopte una perspectiva excesivamente descriptiva, mostrando un cierto temor a expresar sus opiniones o impresiones sobre lo observado.

Por último, cabe plantear aquí un conjunto de reglas para la confección del diario que van a facilitar considerablemente las tareas de análisis de la información en él recogida. Estas reglas son:

— Indicar la fecha y duración de cada sesión de observación.

— Presentar cada sesión de observación con un título que resuma o condense lo más significativo de esa sesión. Esto permitirá encontrar fácilmente el registro de cada sesión, ya que el título será más fácil de recordar que la fecha concreta en que sucedieron los eventos.

— Describir el contexto en el que se realiza la observación y los sujetos que participaron en la misma. Para ello resulta de gran utilidad hacer dibujos o croquis que, a un golpe de vista, describan la distribución espacial de los sujetos.

— Dejar en todo momento márgenes muy amplios (bien a derecha e izquierda, bien sólo a uno de los lados, según el gusto del observador). Es un error común de quienes usan el diario por primera vez no dejar estos márgenes y cuando la información sea analizada, será preciso disponer de ellos para escribir comentarios al margen o indicar dónde empieza y acaba cada categoría de análisis.

— Es preferible que en cada párrafo sólo se aborde un tema. Esto también será de gran ayuda en el momento de la categorización.

3.3. Los sistemas tecnológicos en la observación sistematizada y no sistematizada

Al hablar de los sistemas tecnológicos, nos referimos a las grabaciones en audio y en vídeo y a las fotografías y diapositivas que se utilizan como medios de apoyo complementarios en los procesos de observación. Autores como EVERTSON y GREEN (1989) incluyen los sistemas tecnológicos como una modalidad de registro observacional diferente a los sistemas categoriales, descriptivos y narrativos. Nosotros pensamos que este tipo de sistemas no proporcionan sino los datos en bruto que deberán someterse a observación y registro con posterioridad al momento en que se realiza la grabación o la fotografía. Por eso lo hemos incluido en un apartado diferente.

La principal característica que les atribuimos a los medios tecnológicos como recursos de apoyo a la observación es su gran flexibilidad. Esta flexibilidad los hace extremadamente adaptables a cualquier tipo de registro observacional. Por ello, son complementarios a cualquiera de los registros que hemos descrito con anterioridad. Así, tanto en observaciones con sistemas categoriales como en observaciones no sistematizadas, grabar en vídeo o en magnetofón o tomar fotografías es una forma de registrar de forma permanente los eventos y conductas para volver a ellos siempre que sea necesario y conseguir así un registro más exhaustivo y detallado.

Por lo tanto, una de sus principales ventajas reside en el hecho de que permiten salvar el carácter temporal de los fenómenos facilitando un registro permanente de los mismos. Sin embargo, como decíamos anteriormente, la información en ellos registrada no es sino los datos en bruto, sobre los cuales el observador tendrá que poner en práctica algunos de los tipos de registros vistos con anterioridad.

Otra ventaja de los medios tecnológicos reside en que permiten captar la realidad objeto de observación desde un plano y no desde un punto (EVERTSON y GREEN, 1989). Para entender mejor esta idea, piénsese en un observador que observa y registra una determinada situación. Dadas sus limitaciones humanas, no puede mirar a dos sitios a la vez, luego es posible que si está atento a lo que hace un sujeto, se le escapen otros eventos que no están en su punto de visión. Cuando se graba en vídeo o en audio, esto no sucede. En la cinta se registra todo lo que sucede dentro del plano de visión o de sonido —o de ambas cosas a la vez—. Rebobinando se pueden observar las conductas que han pasado desapercibidas en la situación de observación.

Su uso y aplicación en los procesos observacionales está sometido a las mismas reglas generales de toda observación: seleccionar los tiempos, las situaciones, los sujetos... Sin embargo, de forma adicional hay que adoptar decisiones de carácter técnico relativas a

dónde se va a colocar el aparato, qué tipo de planos se van a tomar, cuál debe ser la iluminación, etc. Si algunas de estas decisiones no se toman convenientemente, corremos el riesgo de que la grabación no se pueda aprovechar.

Entre las dificultades que el uso de medios tecnológicos en la observación puede conllevar, podemos destacar que son recursos relativamente costosos. Por un lado, por la necesidad de disponer del aparato en sí mismo y, por otro, porque en ocasiones es preciso contar con una persona que se dedique a realizar la grabación o la toma de fotografías.

Sin embargo, la crítica principal que ha recibido el uso de medios tecnológicos en la observación, se refiere a que su introducción en los contextos de observación origina una distorsión en las condiciones naturales de los mismos y en el comportamiento de los sujetos observados. Como veremos en el próximo apartado, toda observación puede provocar reacciones en el comportamiento de los sujetos, por lo que la distorsión de la realidad no es achacable en exclusividad a los medios tecnológicos. Para contrarrestar los efectos que su introducción en un determinado contexto puede generar, conviene establecer un período previo de prueba en el que los sujetos se acostumbren a la cámara o la grabadora. Durante estas primeras sesiones en las que el comportamiento puede estar alterado no es necesario grabar, sino hacer como si se grabara. Tras las primeras experiencias con el medio, el contacto cotidiano con el mismo llevará a que los sujetos se habitúen a él y terminen comportándose con naturalidad. Tal y como expresa una cita de GARCÍA FERRANDO y SANMARTÍN (1986), los sujetos no pueden estar alterando su conducta durante un plazo largo de tiempo; acaban acostumbrándose al aparato y decayendo progresivamente su interés.

De todas formas, conviene minimizar el impacto de los medios en los contextos de observación. Si éstos son grandes y aparatosos, es más difícil olvidarse de ellos. Pensemos en un sujeto que tiene la cámara delante de su cara y un micrófono pegado a la boca. Difícilmente podrá olvidarse de su presencia y actuar con naturalidad. Por ello es preferible que los aparatos que se utilicen sean lo más pequeños posible, y se coloquen además en un lugar donde su presencia no resulte intrusiva para las personas a las que se observa.

4. GARANTÍAS CIENTÍFICAS DE LAS TÉCNICAS Y REGISTROS OBSERVACIONALES

Todo instrumento de diagnóstico debe garantizar ciertos indicadores de calidad a la hora de recoger información. Si los instrumentos y técnicas diagnósticos no reunieran estos criterios sería incorrecto formular juicios valorativos y tomar decisiones sobre la base de la información que proporcionan.

Los criterios que habitualmente se toman como referencia para estimar la calidad de la medida que proporciona un instrumento son la fiabilidad y la validez. No obstante, las técnicas observacionales no siempre se diseñan y usan con la finalidad de realizar «mediciones», por lo que estos criterios —validez y fiabilidad— no siempre son aplicables a estas técnicas y los datos que proporcionan.

Por esta razón, en este apartado vamos a hacer una breve exposición de los aspectos a tener en cuenta a la hora de valorar la calidad científica y técnica de los procedimientos observacionales. Ha de tenerse en cuenta que no vamos a referirnos de forma genérica a la validez o fiabilidad de los instrumentos diagnósticos, sino a estos y otros conceptos aplicados específicamente a los procedimientos observacionales. Además, puesto que los tradicionales criterios de validez y fiabilidad son difícilmente aplicables a los registros de observación no sistematizada, también reflexionaremos sobre criterios alternativos que se vienen utilizando en la valoración de las técnicas, procesos y resultados de investigación cualitativa, criterios que serán también de utilidad para otras técnicas que abordaremos en capítulos posteriores. Empezaremos, no obstante, delimitando los principales sesgos que afectan a la observación.

4.1. Sesgos que afectan a las técnicas y registros observacionales

Los sesgos que afectan a la observación pueden agruparse, según IBÁÑEZ (1993), en los tres tipos siguientes:

1) *Sesgos procedentes del sujeto observado*. Se deben a lo que se denomina *reactividad de la observación*, o tendencia a que los sujetos modifiquen su conducta por saberse observados. La influencia de esta fuente de sesgos puede limitarse gracias a procedimientos como: utilizar observadores participantes (es decir, integrados en el contexto en que se observa), solicitar explícitamente al sujeto que no modifique su conducta y fijar un período previo de habituación en el que las conductas registradas no se consideran hasta tanto no apreciamos que el sujeto se está acostumbrando a ser observado. También se pueden utilizar dispositivos ocultos a la vista del sujeto, aunque esta decisión dependerá de la situación, a fin de

garantizar los mínimos éticos necesarios en todo diagnóstico —recogidos, por otra parte, en los códigos deontológicos al uso—.

2) Sesgos procedentes de la persona que observa. Aquí son muchos los factores que pueden afectar negativamente a la observación que realizamos. Algunos de ellos son: el hecho de que la participación del observador afecta a su registro en tanto éste sea subjetivo; la influencia de las expectativas sobre lo que observa (*efecto Rosenthal*); las características generales de la persona que observa (sexo, edad...) cuyos efectos son realmente difíciles de corregir; la falta de entrenamiento; y la habituación, similar al efecto de halo, o tendencia a reproducir las pautas que ha observado en los primeros registros.

3) Sesgos procedentes del sistema de observación. Que englobarían tanto a los sesgos que provienen del instrumento (por ejemplo, el sistema de categorías utilizado), como a las decisiones tomadas en el muestreo. FERNÁNDEZ BALLESTEROS (1992) propone un conjunto de aspectos que disminuyen el riesgo de este tipo de sesgos: 1) claridad de las definiciones conductuales; 2) un reducido número de categorías, ya que en tanto se incrementa el número de las mismas, decrecen la fiabilidad y validez de un sistema; 3) el uso de procedimientos que ya cuentan con garantías científicas; 4) que el observador tenga claro qué tiene que observar cuando el procedimiento que se utiliza no es un sistema categorial previamente elaborado; y 5) que, cuando se empleen escalas de estimación o procedimientos similares, el observador conozca suficientemente al sujeto sobre el que va a realizar el registro.

Otra posible clasificación de fuentes de sesgos en la observación es la que proporciona ANGUERA (1995a), en la que aparecen nuevos aspectos:

— La reactividad, que consiste en la alteración de la naturaleza espontánea de las conductas de los sujetos observados que se ocasiona como consecuencia de que éstos perciben que están siendo objeto de la observación.

— La reactividad recíproca, que tiene lugar cuando el sesgo de reactividad afecta también al observador, el cual se ve influenciado por conocer que el sujeto al que observa no actúa espontáneamente como consecuencia de la situación de observación.

— La autorreactividad que, en este caso, afecta a la auto-observación y, aunque no deja de ser un sesgo, tiene efectos positivos. Implica que el autorregistro ejerce influencia sobre la ocurrencia de la conducta (por ejemplo, una persona que, como parte de una dieta de adelgazamiento, debe anotar cada vez que pica entre horas; es posible que el mero hecho de registrar su conducta le lleve a tomar conciencia de la misma y reduzca su frecuencia).

— La expectativa o influencia de los conocimientos previos del observador o de sus propios deseos y valores a la hora de registrar la conducta.

A la hora de hacer una observación, hemos de ser conscientes de la influencia de estos sesgos, a fin de minimizar su impacto e incidencia en los registros.

4.2. Fiabilidad y validez de los registros observacionales

No existe un consenso generalizado entre los distintos autores a la hora de explicitar cuáles son los criterios y garantías de fiabilidad y validez que la observación debe satisfacer.

GALTUNG (1967[4], cit. por GARCÍA FERRANDO y SANMARTÍN, 1986) señala tres principios que ha de satisfacer toda observación científica: a) el *principio de la intrasubjetividad* o fiabilidad, que hace referencia a que la observación repetida de las mismas respuestas por el mismo observador debe producir los mismos datos; b) el *principio de la intersubjetividad*, que se refiere a que la observación repetida de las mismas respuestas por parte de diferentes observadores debe producir los mismos datos; y c) el *principio de la validez*, que destaca el hecho de que los datos deben obtenerse de tal manera que se puedan realizar inferencias legítimas desde el nivel manifiesto al nivel latente.

QUERA y BEHAR (1997), por su parte, proponen tres criterios para el control de la calidad de las observaciones: a) fiabilidad, entendida como el grado de concordancia entre los datos proporcionados por dos o más observadores cuando registran independientemente la misma conducta en cada ocasión (*fiabilidad interobservadores*); o bien, como la concordancia entre dos registros de una misma sesión obtenidos por el mismo observador en dos momentos diferentes, gracias a la grabación en vídeo (*fiabilidad intraobservador* o consistencia del observador); b) precisión o *grado de concordancia* entre un observador y un criterio establecido o registro estándar; y c) *validez*, o relación existente entre lo que se evalúa y lo que se desea evaluar.

Como se puede observar, los autores citados no sólo hablan de criterios distintos, sino que llaman de distinta forma a la misma cosa, como es el caso de la fiabilidad interobservadores en QUERA y BEHAR (1997) que, en la aportación de GALTUNG (1967) no se considera como fiabilidad, sino como principio de intersubjetividad. Otros autores (BAKEMAN y GOTTMAN, 1989; FOSTER y CONE, 1986, IBÁÑEZ, 1993) añaden a los tradicionales criterios de fiabilidad y validez, la exactitud como una condición más para garantizar la bondad de los datos registrados a partir de la observación. Así, exactitud sería sinónimo del grado en que una medida representa fielmente los rasgos de la conducta objeto de observación. Nosotros consideramos que este concepto es redundante en tanto que está ya recogido en los de validez y fiabilidad.

En relación con la **fiabilidad de los registros observacionales,** nos parece más clara la posición de ANGUERA (1988), quien diferencia entre las siguientes acepciones del término:

— La acepción más común, que sería el *acuerdo entre observadores* que observan simultáneamente una realidad.

— La de estabilidad, en tanto que un mismo observador observe en momentos diferentes.

— Y, finalmente, la acepción que hace referencia a distintos observadores que observan en momentos diferentes.

De todas estas acepciones, es la fiabilidad interobservadores o interjueces la que ha recibido mayor atención. La fiabilidad interjueces hace referencia al grado en que dos observadores independientes, observando las mismas conductas con el mismo instrumento de observación, coinciden o no en sus registros. Aunque es una forma de estimar la bondad de los registros observacionales, hay que tener en cuenta que, en principio, la concordancia entre los observadores no justifica por sí misma que la conducta haya ocurrido, ni que se haya manifestado en los términos apuntados por los observadores, ni que, en un futuro hipotético, esos observadores vayan a categorizar conductas iguales de la misma forma que lo hicieron en el pasado.

Sin embargo, esta forma de entender la fiabilidad presenta como ventaja que ha sido operativizada en diversas fórmulas que permiten obtener una indicación numérica del nivel de concordancia entre observaciones. Algunos de los procedimientos más conocidos para calcular la fiabilidad interjueces son el cálculo del porcentaje de acuerdos y el índice Kappa, ambos métodos son adecuados para sistemas de registro de intervalo. Para cuando los datos están recogidos de forma que es difícil una comparación intervalo a intervalo (por ejemplo, en las escalas de estimación), se usan los denominados métodos para el total de la sesión. En ellos se trata de estimar el acuerdo total de una sesión, mediante procedimientos como la correlación de Pearson entre los registros y el análisis de la varianza.

Por otra parte, entrando ya en la cuestión de la **validez de los registros observacionales,** puede decirse que ésta se puede entender como,

— Relevancia y significatividad de las conductas observadas (validez de contenido).

— Relación existente entre lo que se evalúa y lo que se desea evaluar (validez de constructo).

— Coherencia y concordancia de los datos obtenidos a través de los instrumentos observacionales (validez de criterio).

En los sistemas categoriales (sistemas de categorías, listas de control y escalas de estimación), se puede analizar la validez de contenido de los instrumentos diseñados solicitando la opinión de un grupo de expertos (en la materia o tema que se pretende observar a través del instrumento) sobre el grado en que las manifestaciones relevantes están incluidas en el mismo (PÉREZ JUSTE, 1989). Este autor también propone hacer un análisis de la coherencia y concordancia de los datos obtenidos a través de estos

instrumentos (validez relativa a un criterio) pidiendo a un conjunto de expertos que valoren los comportamientos a observar en determinados sujetos, contrastando después su ordenación con la derivada o resultante de aplicar tales instrumentos. Técnicas estadísticas como los coeficientes de correlación pueden ofrecernos índices empíricos de esta concordancia. Lo que resulta evidente es que, en ambos casos (validez de contenido y relativa a un criterio), el equipo de expertos finalmente seleccionados debe estar ampliamente cualificado para emitir los juicios que se solicitan.

Asimismo, para estudiar la validez de constructo, sería necesario aplicar el instrumento a sujetos distintos en numerosas ocasiones y, a través de las técnicas estadísticas habituales (estudio de las intercorrelaciones, análisis factorial...), observar si el comportamiento de los registros concuerda con la estructura teórica desde la cual está diseñado el instrumento. Sin embargo, esta última faceta de la validez es mucho más difícil de estimar en los procedimientos observacionales. Al igual que las anteriores, sólo podría aplicarse en los sistemas categoriales, donde se parte de una teoría o modelo explicativo a partir del cual se seleccionan las categorías a observar. Por otra parte, su estimación requeriría que el instrumento haya sido aplicado en condiciones iguales a numerosos sujetos, a fin de reunir los mínimos exigidos para la aplicación de ciertas pruebas estadísticas. Las matrices multirrasgo-multimétodo pueden ser un procedimiento de gran utilidad para estimar la validez de constructo de los registros observacionales, en tanto permiten comparar varios métodos de medición de uno o varios constructos.

4.3. Criterios de calidad alternativos para los registros de observación no sistematizados

Los criterios de fiabilidad y validez abordados en los párrafos precedentes son difícilmente aplicables a estrategias de observación mediante las cuales la información que se recaba es cualitativa y no susceptible de ser transformada en códigos numéricos. Por otra parte, ante la inexistencia de un conjunto de categorías deductivas a observar que unifique la perspectiva de distintos observadores, resulta difícil estimar si se producen o no acuerdos —en términos numéricos de porcentaje de acuerdo a como se ha expuesto con anterioridad—.

Por esta razón, para la investigación cualitativa en general, y sus diferentes procedimientos de recogida de información en particular, se han venido utilizando criterios alternativos a través de los cuales se pretende paliar la incidencia de sesgos que distorsionen la información.

Autoras como GOETZ y LE COMPTE (1988) hacen un esfuerzo por adaptar a la investigación cualitativa los tradicionales criterios de validez y fiabilidad. Sin embargo, otros autores, como GUBA (1983) y LINCOLN y GUBA (1985), partiendo de las características diferenciales de la investigación cualitativa, aportan criterios distintos que

responden a la lógica propia de este tipo de investigación. A continuación, revisamos los criterios, ya clásicos, propuestos por estos autores.

Frente al valor de verdad de los enfoques cuantitativos, LINCOLN y GUBA (1985) proponen el **criterio de credibilidad** que hace referencia al isomorfismo entre los datos recogidos por el investigador y la realidad. Distintos procedimientos contribuyen a garantizar este criterio:

— La *observación persistente o prolongada*. Que permite que el observador pueda comprobar los prejuicios propios y de los participantes, así como un enfoque más intenso de los aspectos más característicos de la situación, lo que aporta mayor verosimilitud.

— La *triangulación*. Es un procedimiento muy utilizado en la investigación cualitativa, a través del cual se compara y somete a contraste la información recogida sobre una situación desde distintos ángulos y perspectivas. La triangulación puede hacerse de diferentes formas (COHEN y MANION, 1990): a) triangulación de los datos en distintos momentos temporales y situacionales; b) triangulación de investigadores con distintos puntos de vista; c) triangulación teórica, que implica analizar un fenómeno desde distintos modelos y teorías; d) triangulación metodológica, que consiste en utilizar varias estrategias y métodos; e) triangulación de sujetos, a través de la cual la información se compara según el punto de vista de personas que desempeñan roles distintos o que enfocan de forma visiblemente distinta la situación; f) triangulación múltiple, en la que se combinan varios tipos de triangulación y niveles de análisis.

— El *intercambio de opiniones* con otros investigadores.

— *Recogida de material referencial*. Se refiere a recoger información mediante vídeos, documentos o grabaciones como referencias a partir de las cuales se puede hacer una comparación.

— *Comprobaciones con los participantes*. En las observaciones no sistematizadas, a fin de evitar errores y conseguir más información sobre los puntos oscuros, es conveniente llevar a cabo revisiones y comprobaciones con las personas que son observadas, las cuales pueden corroborar si se ha captado o no el significado de la situación.

Por otra parte, GUBA (1983) y LINCOLN y GUBA (1985) proponen el **criterio de transferibilidad** como alternativo al de generalizabilidad o aplicabilidad usual en los enfoques cuantitativos. La transferibilidad hace referencia al grado en que pueden aplicarse los resultados de una investigación a otros contextos y situaciones. Las estrategias que lo posibilitan son el muestreo teórico (situaciones y contextos múltiples), las descripciones copiosas (que permiten establecer comparaciones y hacer más extensibles las generalizaciones) y la recogida abundante de información (para comprobar el grado de correspondencia con otros contextos).

En correspondencia con el criterio de consistencia, los autores citados proponen el de **dependencia** que alude a la estabilidad de los datos y registros, y a la posibilidad de obtener los mismos resultados cuando se replica el estudio con los mismos o similares sujetos y contextos. Las estrategias a utilizar para garantizar este criterio son: establecer pistas de revisión y la réplica paso a paso (para revisar los procedimientos seguidos), la auditoría de dependencia (revisión del proceso de investigación por investigadores externos), y los métodos solapados (utilización de varios métodos y procedimientos de recogida de información que permitan perspectivas diferentes en el análisis de un sujeto o situación; por ejemplo, comparando los registros de observación con grabaciones en vídeo).

Por último, el **criterio de confirmabilidad** de GUBA (1983) y LINCOLN y GUBA (1985) se corresponde con el concepto tradicional de objetividad. Este criterio permite garantizar que los resultados de una investigación no están sesgados por motivaciones, intereses y perspectivas del investigador. Consiste, por tanto, en confirmar la información y la interpretación de los resultados. Las estrategias que contribuyen a la confirmabilidad son: la auditoría de confirmabilidad (un auditor externo comprueba la correspondencia entre los datos y las inferencias o interpretaciones), los descriptores de baja inferencia (registros del fenómeno lo más precisos posibles, como las transcripciones textuales) y los ejercicios de reflexión.

Estos criterios siguen teniendo una gran aceptación a la hora de valorar las investigaciones cualitativas, sin embargo, más recientemente otros autores han propuesto nuevos indicadores de calidad para los procesos y procedimientos de investigación cualitativa que abordan muchos aspectos no considerados por Guba y Lincoln. El trabajo de SANDÍN (2000) aporta una excelente síntesis sobre el estado actual de esta cuestión.

Referencias bibliográficas

ANDER-EGG, E. (1980). *Técnicas de investigación social*. Buenos Aires: El Cid.

ANGUERA, M. T. (1982). *Metodología de la observación en las ciencias humanas*. Madrid: Cátedra.

ANGUERA, M. T. (1988). *Observación en la escuela*. Barcelona: Graó.

ANGUERA, M. T. (1991). Proceso de categorización. En M.T. Anguera (ed.), *Metodología observacional en la investigación psicológica, Vol. I, Fundamentación* (pp. 115-167). Barcelona: PPU.

ANGUERA, M. T. (1995a). Metodología cualitativa. En M.T. Anguera et al., *Métodos de investigación en psicología* (pp. 513-522). Madrid: Síntesis.

ANGUERA, M. T. (1995b). Tratamiento cualitativo de datos. En M.T. Anguera et al., *Métodos de investigación en psicología* (pp. 549-576). Madrid: Síntesis.

ANGUERA, M. T. (1995c). Recogida de datos cualitativos. En M.T. Anguera et al., *Métodos de investigación en psicología* (pp. 523-547). Madrid: Síntesis.

AVIA, M. A. (1983). La auto-observación. En R. Fernández-Ballesteros y J.A. Carrobles (eds.), *Evaluación conductal: metodología y aplicaciones*. Madrid: Pirámide.

BAKEMAN, R. y GOTTMAN, J. R. (1989). *Observación de la interacción: introducción al análisis secuencial.* Madrid: Morata.

BARDIN, L. (1986). *Análisis de contenido*. Madrid: Akal.

BEHAR, J. (1991). Observación y análisis de la producción verbal. En M.T. Anguera (ed.), *Metodología observacional en la investigación psicológica, Vol. I, Fundamentación* (pp. 330-389). Barcelona: PPU.

BLANCO, A. y ANGUERA, M. T. (1991). Sistemas de codificación. En M.T. Anguera (ed.), *Metodología observacional en la investigación psicológica, Vol. I, Fundamentación* (pp. 193-239). Barcelona: PPU.

BLANCO, F. (1994). *La evaluación en la educación secundaria*. Salamanca: Amarú Ediciones.

CABRERA, F. A. y ESPÍN, J. V. (1986). *Medición y evaluación educativa*. Barcelona: PPU.

CARRERAS, M. V. (1991). Métrica del registro observacional. En M.T. Anguera (ed.), *Metodología observacional en la investigación psicológica, Vol. I, Fundamentación* (pp. 169-192). Barcelona: PPU.

COHEN L. y MANION, L. (1990). *Métodos de investigación educativa*. Madrid: La Muralla.

COLÁS, P. (1994). Los métodos descriptivos. En P. Colás y L. Buendía, *Investigación educativa* (pp. 177-200). Sevilla: Alfar.

DENDALUCE, I. (1988). *Aspectos metodológicos de la investigación educativa*. Madrid: Narcea.

EVERTSON, C. y GREEN, J. (1989). La observación como indagación y método. En M. Wittrock (coord.), *La investigación en la enseñanza*, *tomo II* (pp. 303-421). Barcelona: Paidós.

FERNÁNDEZ BALLESTEROS, R. (1992). La observación. En R. Fernández Ballesteros (coord.), *Introducción a la evaluación psicológica* (pp.137-182). Madrid: Pirámide.

FOSTER, S. L. y CONE, J. D. (1986). Design and use of direct observation procedures. En A. Ciminero, K.S. Calhoum y H.E. Adams (eds.), *Handbook of behavioral assessment*. New York: Wiley.

GARCÍA FERRANDO, M. y SANMARTÍN, R. (1986). La observación científica y la obtención de datos sociológicos. En M. García Ferrando, J. Ibáñez y F. Alvira (comp.), *El análisis de la realidad social. Métodos y técnicas de investigación* (pp. 95-122). Madrid: Alianza Editorial.

GARCÍA JIMÉNEZ, E. (1997). *La recogida de información*. Material de la asignatura Diagnóstico Pedagógico, inédito. Sevilla: Facultad de Ciencias de la Educación.

GOETZ, J. P. y LECOMPTE, M. D. (1988). *Etnografía y diseño cualitativo en investigación educativa*. Madrid: Morata.

GUBA, E. G. (1983). Criterios de credibilidad en la investigación naturalista. En J. Gimeno y A. Pérez (ed.), *La enseñanza: su teoría y su práctica* (pp. 148-165). Madrid: Akal.

HERBERT, J. D. y NELSON-GRAY, R. O. (1997). La auto-observación. En G. Buela-Casal y J.C. Sierra (dirs.), *Manual de evaluación psicológica* (pp. 343-367). Madrid: Siglo Veintiuno Editores.

HERNÁNDEZ DE LA TORRE, M. E. (1997). La observación en el contexto de la evaluación psicopedagógica de las necesidades educativas. En H. Salmerón (coord.), *Diagnosticar en educación* (pp.417-424). Granada: FETE-UGT.

IBÁÑEZ, I. (1993). La observación. En W. Peñate, P. Mataud e I. Ibáñez, *Evaluación psicológica: Concepto y técnicas de análisis* (pp. 223-260). Valencia: Promolibro.

KERLINGER, F.N. (1975). *Investigación del comportamiento*. México: Nueva Editorial Interamericana.

LINCOLN, Y. y GUBA, E. (1985). *Naturalistic enquiry*. Beverly Hills, CA: Sage.

LÓPEZ GÓRRIZ, I. (1997). El diario: un instrumento para diagnosticar procesos grupales. En H. Salmerón (coord.), *Diagnosticar en educación* (pp.433-439). Granada: FETE-UGT.

LÓPEZ, I. y otros (1996). La formación en el proceso de formación del profesorado en educación de adultos desde un proyecto de investigación-acción. *Cuestiones Pedagógicas, 12*, 209-222.

LÓPEZ, I. y PADILLA, M. T. (1997). La formación del pedagogo a través de las prácticas: reflexiones a la luz de la experiencia. *Cuestiones Pedagógicas, 13*, 227-235.

LOSADA, J. L. (1999). *Metodología observacional. Cuaderno de problemas*. La Coruña: Penta.

MUCCHIELLI, R. (1974). *El cuestionario en la encuesta psico-social* (4ª edición). Madrid: Ibérico Europea de Ediciones.

PÉREZ JUSTE, R. (1989). Los datos para la evaluación. En R. Pérez Juste y J.M. García Ramos, *Diagnóstico, evaluación y toma de decisiones* (pp. 73-193). Madrid: Rialp.

PÉREZ JUSTE, R. (1990). Recogida de información en el diagnóstico pedagógico. *Bordón, 42*(1), 17-29.

PÉREZ JUSTE, R. (1991). *Pedagogía experimental. La medida en educación*. Madrid: UNED.

QUERA, V. (1991). Muestreo y registro observacional. En M.T. Anguera (ed.), *Metodología observacional en la investigación psicológica, Vol. I, Fundamentación* (pp. 241-329). Barcelona: PPU.

QUERA, V. y BEHAR, J. (1997). La observación. En G. Buela-Casal y J.C. Sierra (dirs.), *Manual de evaluación psicológica* (pp. 315-341). Madrid: Siglo Veintiuno Editores.

SANDÍN, M.P. (2000). Criterios de validez en la investigación cualitativa: de la objetividad a la solidaridad. *Revista de Investigación Educativa, 18* (1), 223-242.

SIERRA BRAVO, R. (1983). *Técnicas de investigación social*. Madrid: Paraninfo.

SPRADLEY, J. P. (1980). *Participant observation*. New York: Holt, Rinehart y Winston.

ZABALZA, M., MONTERO, L. y ÁLVAREZ, Q. (1986). Los diarios de los alumnos de Magisterio en prácticas como instrumento de formación profesional. En L.M. Villar (ed.), *Actas del Primer Congreso Internacional sobre Pensamientos de los Profesores y Toma de Decisiones*. Sevilla: Servicio de Publicaciones de la Universidad de Sevilla.

CAPÍTULO

3

INTRODUCCIÓN AL DISEÑO Y APLICACIÓN DE LAS TÉCNICAS E INSTRUMENTOS DE ENCUESTA

1. LA ENCUESTA COMO ESTRATEGIA DIAGNÓSTICA

No siempre resulta posible acceder de forma directa y «sin intermediarios» a la información necesaria para un diagnóstico pedagógico. Determinadas conductas y situaciones tienen lugar fuera del espectro perceptivo de la persona que realiza el diagnóstico. Así, explorar a través de la observación las percepciones que una persona tiene de una situación o problema, sus creencias, sentimientos y planes resulta muy difícil y poco efectivo. Pensemos también en las conductas o fenómenos que sucedieron en el pasado, o en la conducta íntima o privada (la actividad sexual de una persona es un buen ejemplo de ello). Ante este tipo de eventos la observación se plantea como una tarea imposible o ineficaz, por lo que el profesional del diagnóstico se ve obligado a recurrir a otros medios diferentes a fin de conocer aquello en lo que tiene interés. Uno de estos medios posibles consiste en preguntar al sujeto o sujetos que tiene/n la información buscada. Así, cuando solicitamos información desde el punto de vista de otras personas a través de la formulación de preguntas, estamos haciendo uso de los métodos de encuesta.

Hecha esta presentación, debe quedar claro que, desde nuestro punto de vista, la peculiaridad que define a las técnicas de encuesta —bien se use el cuestionario o la entrevista— es que la información se obtiene mediante la formulación de preguntas. En opinión de SELLTIZ et al. (1976), preguntar es la estrategia útil cuando se quiere obtener información sobre *qué sabe*, *cree*, *espera*, *siente* o *quiere*, *intenta hacer* o *ha hecho* una persona, y acerca de sus *explicaciones* o *razones* para cualquiera de los puntos señalados.

Al igual que comentábamos en el capítulo dedicado a la observación, no está claro para todos los autores qué estrategias concretas se deben englobar bajo el término de «encuesta», aunque a diferencia de la observación, sí resulta posible encontrar un mayor nivel de consenso en la inclusión, dentro de las técnicas de encuesta, de la entrevista —en sus diferentes modalidades— y el cuestionario. Donde aparecen mayores dificultades a la hora de utilizar esta clasificación es en los diferentes procedimientos que se incluyen dentro del epígrafe «cuestionario». Así, encontramos autores como MUCCHIELLI (1974) que considera que instrumentos como las escalas de actitudes y las técnicas proyectivas también son cuestionarios.

Ahora bien, en las técnicas de encuesta hay dos formas de interacción distintas entre encuestador-encuestado: la interacción personal en que el investigador y el sujeto se encuentran cara a cara y el primero plantea las preguntas al segundo (mediante la entrevista), y la interacción impersonal, en la que el investigador está «representado» en el cuestionario que ha hecho llegar al sujeto, pero no está presente físicamente hablando. FOX (1981) también habla de una «interacción mixta» en la que el cuestionario se aplica

a grupos de sujetos con una previa indicación oral de las instrucciones y aclaración de dudas.

Potencialmente, las técnicas que se basan en el método de encuesta presentan como característica su gran versatilidad, ya que pueden utilizarse prácticamente en cualquier campo (actitudes, opiniones, hábitos...), y prácticamente con cualquier grupo humano (GARCÍA FERRANDO, 1986), lo que, en palabras de MANZANO, ROJAS y FERNÁNDEZ (1996:20), hace de ellas «un verdadero todo-terreno».

2. LA ENTREVISTA EN EL DIAGNÓSTICO Y LA EVALUACIÓN

La entrevista es una estrategia que puede servir a múltiples fines. Puede usarse como medio para recabar información de grandes poblaciones a fin de probar hipótesis científicas, para seleccionar a los mejores candidatos a un puesto de trabajo, como medio para conseguir cambios en la conducta de un sujeto, para conocer las opiniones de consumidores y usuarios potenciales de un servicio o producto, etc.

Por otra parte, la entrevista, en sus diversas modalidades, es considerada una estrategia fundamental en los procesos diagnósticos y orientadores. No sólo sirve al fin de recoger información en torno a una situación o problema, sino que constituye también una estrategia de intervención. Como plantea GARCÍA MARCOS (1983) —aunque refiriéndose estrictamente a la evaluación psicológica—, la entrevista no sólo es el instrumento de evaluación que precede a cualquier modalidad de intervención o toma de decisiones, sino que es el sistema de interacción mediante el cual se realiza todo el proceso de evaluación-intervención. Por ello, no sólo cumple una función diagnóstica, sino que también sirve a otras funciones, como son la recogida amplia de información sobre el sujeto (historia, trayectoria laboral y profesional, situación familiar, etc.), la función motivadora (establecimiento de una relación positiva, cambio de actitudes y expectativas, etc.) y la función terapéutica (puesta en práctica de estrategias de intervención).

Lo que caracteriza a la entrevista, frente a otros procedimientos de encuesta, es la *interacción personal* que en ella tiene lugar. Así, la entrevista es un tipo de conversación en la que dos personas, o más de dos, interactúan de forma verbal y no verbal con el propósito de cumplir un objetivo previamente definido (MATARAZZO, 1965[5], cit. por GARAIGORDOBIL, 1998). O, de forma más simple, en la definición tradicional de BINGHAM y MOORE (1924, cit. por MATEU y FUENTES, 1992), una *conversación con un propósito*.

Algunos autores (GARAIGORDOBIL, 1998; GARCÍA MARCOS, 1983; PEÑATE, 1993; SILVA 1992) señalan un conjunto de aspectos comunes en las definiciones de entrevista:

— Constituye una relación entre dos o más personas.

— Es una vía de comunicación simbólica bidireccional, preferentemente oral.

— Requiere unos objetivos prefijados y conocidos, así como una asignación de roles que significa un control de la situación por parte del entrevistador. Es por ello que en la entrevista tiene lugar una *relación interpersonal asimétrica*.

Por su parte, ROSSELL (1989) señala tres características en la entrevista:

— Es una relación, lo que implica que la entrevista tiene que ser algo más que la aplicación mecánica de una técnica.

— Es una técnica, por lo que en ella se operativizan contenidos teóricos con la finalidad de estudiar situaciones.

— Es un proceso con diferentes fases.

En todo caso, aunque la entrevista se desarrolla en el marco de una relación entre dos o más personas y a través de la comunicación oral, conviene no confundirla con una simple conversación. No sólo está *orientada hacia unos objetivos*, sino que en ella uno de los participantes es el que ejerce el control y la dirige hacia una finalidad dada.

MATEU y FUENTES (1992) y SILVA (1992) señalan las siguientes ventajas en la entrevista como instrumento de evaluación: la relación interpersonal que implica; la flexibilidad que permite al entrevistador de adaptarse sobre la marcha a las peculiaridades del entrevistado; la posibilidad de observación del comportamiento del entrevistado, tanto en la dimensión verbal como no verbal; la posibilidad de registrar grandes cantidades de información de carácter subjetivo (emociones, por ejemplo) y de tipo biográfico (trayectoria vital, etc.); y la posibilidad de evaluar a personas que no podrían ser examinadas con otros instrumentos (por ejemplo, las que no saben leer y escribir).

Entre sus limitaciones, estos autores señalan el coste relativamente elevado de tiempo y esfuerzo que requiere por parte del entrevistador y la interferencia de sesgos de variada procedencia (del entrevistador, del entrevistado, de la relación y de la situación), que mencionaremos en el apartado final de este capítulo.

A la hora de clasificar las entrevistas es posible utilizar diversos criterios de clasificación. GARAIGORDOBIL (1998) y GARCÍA MARCOS (1983) diferencian entre tres formas diferentes de clasificación:

— *Según los objetivos*, puede diferenciarse entre aquellas entrevistas que persiguen la recogida de información (entrevistas de investigación, denominadas «encuestas» por SILVA, 1992) y las que persiguen la intervención (bien sean de evaluación o terapéuticas). Utilizando este mismo criterio, PEÑATE (1993) clasifica a las entrevistas en: de investigación, diagnósticas y terapéuticas.

— *Según el grado de estructuración*, la cual puede estar referida a las preguntas, las respuestas, la secuencia de preguntas, el registro y la interpretación de la información. Es por ello que no puede hablarse de estructuración o no estructuración de forma dicotómica, sino de grados dentro de un continuo (GARAIGORDOBIL, 1998; SILVA, 1992).

— *Según el marco teórico del entrevistador*, podría diferenciarse entre la entrevista fenomenológica, psicodinámica y conductual.

Entre los criterios de clasificación expuestos, el que más se utiliza es el referido al grado de estructuración, criterio que será el que utilizaremos aquí. Nuestra elección no sólo se debe a que este criterio es el de mayor interés en la actualidad, sino también al hecho de que nos permite prestar atención a un aspecto de la entrevista de gran relevancia a la hora de hablar de su preparación, realización e interpretación. Así, según el grado de estructuración, las consideraciones a tener en cuenta para diseñar y aplicar entrevistas van a ser muy diferentes, por lo que aquí nos interesa estudiarlas por separado, dada la finalidad de esta obra de adentrar al lector en el diseño y aplicación de instrumentos diagnósticos.

Desde esta perspectiva —la *estructuración como un continuo*—, cabe diferenciar también un tipo de entrevistas que se situaría en torno al punto intermedio de dicho continuo. Nos referimos a las entrevistas semiestructuradas. Existe una gran variedad en lo que podemos denominar como «entrevista semiestructurada», encontrándose opiniones que, bajo este término, caracterizan a las entrevistas que trabajan con preguntas abiertas cuyo orden y secuencia dentro de la entrevista está previamente fijado. En otros casos, se entiende por entrevista semiestructurada a aquella en la que ni las preguntas ni la secuencia están prefijadas, pero se trabaja de acuerdo con esquemas o pautas que deben ser cumplimentadas (GARAIGORDOBIL, 1998).

Por lo tanto, según su grado de estructuración, las entrevistas pueden ser (PEÑATE, 1993):

— *Entrevistas estructuradas:* en las que el texto y secuencia de las preguntas están prefijados así como, en algunos casos, las respuestas a dar por el entrevistado y la situación de entrevista.

— *Entrevistas semiestructuradas*. Diferencia aquí dos opiniones: quienes consideran que este tipo de entrevistas son las que no tienen prefijadas las preguntas pero sí poseen una pauta o esquema de entrevista, y quienes consideran que las entrevistas semiestructuradas son aquéllas que constan de preguntas abiertas y una secuencia prefijada.

— *Entrevistas no estructuradas*: en las que ni las preguntas están prefijadas ni existen esquemas o indicaciones que marquen el rumbo de la entrevista, aunque la entrevista puede estar guiada por «esquemas implícitos» (esto vendría a referir que, a pesar de la ausencia de normas y reglas que condicionen o limiten la entrevista, el entrevistador siempre posee alguna orientación teórica o una formación en la dirección de este tipo de entrevista y, al menos, parte de unos supuestos sobre los cuales desarrolla la misma).

Es nuestra intención referirnos en este texto a los dos polos del continuo; es decir, por un lado, analizaremos las peculiaridades de las entrevistas que tienen el grado máximo de estructuración (aquellas en las están prefijadas tanto las preguntas, como la respuestas, como la secuencia de preguntas, el tipo de registro, la interpretación...) y, por otro, haremos referencia a un tipo de entrevista que tiene el grado mínimo de

estructuración, las entrevistas abiertas, no estructuradas o en profundidad. Entendemos que las modalidades de entrevista que se sitúan en la zona central (lo que anteriormente hemos denominado como entrevistas semiestructuradas), puesto que pueden responder a un patrón muy diverso, como ya hemos planteado, habrán de analizarse según se aproximen más o menos a uno de los dos extremos del continuo. Por ello, si una entrevista semiestructurada se aproxima más al grado máximo de estructuración, aquellas características planteadas para las entrevistas estructuradas pueden aplicarse también a ellas con los oportunos matices, y viceversa.

En el continuo de estructuración de las entrevistas, el máximo grado de control se sitúa en las estructuradas y el mínimo en las no estructuradas (PEÑATE, 1993). Mientras que la necesidad de entrenamiento del entrevistador se diferenciaría en el sentido contrario: hace falta mayor entrenamiento para desarrollar entrevistas no estructuradas, y muy poco entrenamiento previo para las entrevistas estructuradas.

Asimismo, puede decirse que en las entrevistas estructuradas la fase de preparación o diseño de la entrevista es la más importante. En la medida en que las preguntas estén bien formuladas y seleccionadas, y se hayan previsto todas las posibles contingencias que, durante la entrevista, pueden ocurrir, la aplicación de la entrevista no supondrá ninguna dificultad y no se requerirán entrevistadores altamente cualificados. En el caso de las entrevistas no estructuradas, sucede justamente lo contrario. Si bien en la fase de diseño son escasas las tareas a realizar, salvo decidir a quién entrevistar y qué posibles temáticas indagar, la fase de aplicación y desarrollo de la entrevista es de extrema importancia. Crear el clima adecuado, plantear las preguntas de forma que generen información abundante, reconducir la entrevista, tocar temas que pueden resultar problemáticos, etc. serán aspectos esenciales que requerirán ciertas cualidades y bastante experiencia previa por parte del entrevistador.

Para terminar con esta introducción inicial a los distintos tipos de entrevista, conviene recordar que, en opinión de SILVA (1992), los distintos grados de estructuración de la entrevista sólo aparecen como excluyentes si se les considera en un corte transversal, pero no en el corte longitudinal que un proceso de evaluación usualmente implica. Según este autor, por lo general, está más indicada una menor estructuración en los contactos diagnósticos iniciales, y una mayor estructuración en los contactos diagnósticos posteriores. Asimismo, el continuo inestructuración-estructuración se puede apreciar también en el desarrollo de un esquema determinado de entrevista, en la medida en que éste va mejorando y perfeccionándose. Así, cuanto más conocemos de un fenómeno y de cómo reaccionan los sujetos ante el mismo, más factible resultará estructurar una entrevista inicialmente abierta.

2.1. Entrevistas abiertas o no estructuradas

Si bien existe cierto consenso entre los distintos autores en torno a los objetivos y características de la entrevista estructurada, no puede decirse lo mismo respecto a las entrevistas que están en el polo opuesto, es decir, aquellas que se caracterizan por carecer de una estructura rígida, fijada y acotada con antelación a su realización. Encontramos habitualmente distintas denominaciones respecto a esta modalidad de entrevista, resultando que no siempre se habla de lo mismo. Esta pluralidad de denominaciones no sólo proviene de los diversos campos de aplicación habitual de la entrevista (psicología clínica, investigación social, etc.), sino que, dentro de cada campo concreto, se encuentran diferentes opciones teóricas que dan lugar a diferentes formas concretas de denominar y llevar a la práctica la entrevista abierta. Para ejemplificar esta situación, baste mencionar que autores como COHEN y MANION (1990) diferencian, además de la entrevista estructurada, entre tres tipos de entrevistas que pueden calificarse de abiertas:

— *Entrevista no estructurada*: en la que se parte de una situación abierta de mayor flexibilidad y libertad. Aunque planificada, el entrevistador va tomando decisiones que afectan al contenido y secuencia de las preguntas durante su realización.

— *Entrevista no directiva:* cuyos rasgos característicos son la dirección o control mínimos exhibidos por el entrevistador y la libertad que el informante tiene para expresar sus sentimientos subjetivos.

— *Entrevista dirigida:* que se centra en las respuestas subjetivas del informante a una situación conocida. Cuando resulta conveniente, en la entrevista dirigida, el entrevistador representa un papel más activo, introduciendo indicaciones orales más explícitas para activar un informe concreto de respuesta por parte del informante.

Por su parte, ORTÍ (1986) utiliza la denominación de *entrevista abierta o libre* en la cual se pretende profundizar en las motivaciones personalizadas de un caso individual frente a cualquier problema de índole social. Según este autor, este tipo de entrevista consiste en un diálogo *cara a cara*, directo y espontáneo, de una cierta concentración e intensidad entre el entrevistado y un profesional más o menos experimentado, que orienta el discurso lógico y afectivo de la entrevista.

Algunos autores (por ejemplo, WOODS, 1989), usan el término «conversación» y no el de entrevista para caracterizar a las entrevistas abiertas y no estructuradas. De esta forma y, en opinión de este autor, se indica mejor «un proceso libre, abierto, democrático, bidireccional e informal, en el que las personas pueden manifestarse tal como son, sin sentirse atados a papeles predeterminados» (WOODS, 1989:82).

TAYLOR y BOGDAN (1990) prefieren denominar a este tipo de entrevista abierta y no estructurada con el término *entrevistas en profundidad*, a las que definen como:

> «...reiterados encuentros cara a cara entre el investigador y los informantes, encuentros éstos dirigidos hacia la comprensión de las perspectivas que tienen los informantes respecto de sus vidas, experiencias o situaciones, tal como las expresan con sus propias palabras» (TAYLOR y BOGDAN, 1990:101).

Al igual que Woods, estos autores plantean que estas entrevistas siguen el modelo de una conversación entre iguales y no un intercambio formal de preguntas y respuestas. De hecho, en ocasiones, este tipo de entrevistas se consideran como «una forma de observación participante» (WOODS, 1989:104).

Por lo tanto, independientemente de la denominación que utilicemos para referirnos a ellas, entendemos que estamos ante una modalidad de entrevistas que se desarrollan en una situación «abierta», donde hay una mayor flexibilidad y libertad en tanto no existe un esquema rígido de preguntas y alternativas de respuesta que condicionen el desarrollo de la entrevista. La inexistencia de este esquema o estructura previa implica que este tipo de entrevista no es adecuada para contrastar una teoría o un modelo explicativo de un fenómeno o evento determinado. Su finalidad consiste más bien en profundizar en una realidad hasta alcanzar explicaciones convincentes o conocer cómo otras personas —participantes de un contexto o situación educativa— ven y perciben su realidad, siempre desde su propia perspectiva. Afirman ALONSO (1994) y ORTÍ (1986) que la entrevista abierta es muy productiva para el estudio de casos típicos o extremos, en el cual la actitud de ciertos sujetos encarna, con toda su riqueza, el modelo ideal de una determinada actitud, mucho menos cristalizada en la «media» del colectivo de referencia.

A este tipo de entrevistas también se les suele identificar como «entrevistas informales», ya que se realizan en las situaciones y escenarios más diversos: mientras se ayuda al entrevistado a realizar una tarea, en el transcurso de un acontecimiento, en una cafetería... (GARCÍA JIMÉNEZ, 1997). Al mismo tiempo y dado que el entrevistador no adopta un rol inflexible, se consideran informales también por el hecho de que todo es negociable en la entrevista: los entrevistados pueden opinar sobre la conveniencia o no de una pregunta, corregirla, hacer puntuaciones, etc. Esta negociación continua de los papeles y circunstancias de la entrevista, concede al entrevistado una sensación de poder y de responsabilidad personal en el proceso de diagnóstico o evaluación (WOODS, 1989).

Por tanto, a diferencia de las entrevistas estructuradas, en este tipo de entrevistas las reglas están en un proceso de renegociación permanente en el curso mismo del diálogo: frente a un esquema clásico «estímulo-respuesta», donde los papeles están cerrados y no existe retroalimentación, en la entrevista abierta los papeles tienden a estar más abiertos, siendo necesaria una escucha activa y metódica por parte del entrevistador (ALONSO, 1994).

Si bien no existe, como hemos indicado, un esquema rígido al cual tenga que ajustarse la aplicación de la entrevista, en las entrevistas abiertas cabe hablar de *esquemas implícitos.* Así, aunque en ellas el entrevistador adopta una actitud de escucha ante lo que el sujeto contesta, siempre parte de unos esquemas que condicionan el desarrollo de la entrevista. Sus opiniones previas en torno a la cuestión que está diagnosticando le van a llevar a formular unas preguntas y no otras, a profundizar en unos temas y no en otros; su formación previa, la información que ha recabado en entrevistas anteriores, los

modelos teóricos a los que se adscribe... son condicionantes tácitos que ejercen una influencia indirecta en las entrevistas etnográficas.

Por lo tanto, pese a la no estructuración y apertura existente en la entrevista en profundidad, ésta no debe confundirse con una conversación. SPRADLEY (1979) señala tres elementos diferenciadores de la entrevista abierta frente a la mera conversación:

— Por un lado, la *existencia de un propósito explícito.* Así, aunque la entrevista puede comenzar como una charla sobre temas intrascendentes, poco a poco se irá focalizando en un conjunto de temas relevantes para el proceso de diagnóstico que está en marcha.

— Por otro, la presentación de *explicaciones etnográficas*. El entrevistador debe ofrecer explicaciones sobre la finalidad y orientación general del estudio que realiza, sobre los términos o temas que desea aborde el informante, sobre las preguntas que realiza...

— Por último, la formulación de *cuestiones etnográficas*. El tipo de cuestiones que se formulan en la entrevista presentan unas características peculiares que tienden a acentuar el carácter específico y diferencial de este tipo de entrevistas.

A estos rasgos distintivos de la entrevista frente a la conversación libre entre dos o más personas, GARCÍA JIMÉNEZ (1997) añade otros: el turno de palabra es menos equilibrado; repetir reemplaza la regla habitual de evitar la repetición; sólo el entrevistador expresa interés e ignorancia frente a lo que dice el entrevistado; el estímulo continuo al entrevistado para que hable, en lugar de la búsqueda de la brevedad. Todos estos rasgos hacen que, como en toda entrevista, en la entrevista abierta se produzca una *relación interpersonal asimétrica*, lo que implica que existen diferencias entre los roles que adoptan entrevistado y entrevistador.

Por otra parte, como planteamos anteriormente, en este tipo de entrevistas la pericia y formación del entrevistador resulta esencial para obtener la mejor y más amplia información. Dado que el desarrollo de este tipo de entrevistas es más flexible, se requiere más habilidad por parte del entrevistador. Su función, en este tipo de entrevista, es servir de *catalizador* de una expresión exhaustiva de los sentimientos y opiniones de los sujetos y del ambiente de referencia dentro del cual tienen personal significado sus sentimientos y opiniones (SELLTIZ et al., 1976).

Tras lo expuesto anteriormente para caracterizar las entrevistas abiertas o en profundidad, parece claro que son necesarias varias sesiones de encuentro para establecer una relación de confianza que permita la profundización progresiva en los temas que se abordan. Es difícil que en el marco de un intervalo breve de tiempo —media hora o una hora—, se posibilite un proceso tal que conduzca a la generación de la información que con ella se busca, por lo que resulta muy habitual que este tipo de entrevistas tengan una duración más amplia que puede abarcar desde varias horas a

varias sesiones de encuentro, como ya hemos apuntado. Ello dependerá de las circunstancias que condicionen su aplicación.

2.1.1. Preparación de la entrevista no estructurada

Resulta difícil abordar un apartado dedicado a la preparación de las entrevistas abiertas, cuando en este tipo de entrevistas no existe una estructura previa que sea necesario explicitar, a fin de construir el formato final de la entrevista con antelación a su aplicación. De hecho, muchas de las cuestiones que aquí se abordan se refieren a aspectos que hay que tener en cuenta cara al desarrollo de la entrevista, lo que parece confuso teniendo en cuenta que hemos dedicado otro apartado al tema de la aplicación de las entrevistas abiertas.

No obstante, ya hemos planteado que en las entrevistas no estructuradas existen esquemas implícitos que condicionan los temas que se abordan en ellas y, por tanto, las preguntas que se formulan. Por otra parte, nuestra experiencia nos dice que es conveniente realizar algunas tareas relacionadas con su preparación, a fin de evitar que la entrevista se nos vaya de las manos o resulte en un producto final difícilmente diferenciable de lo que se obtiene tras la aplicación de una entrevista estructurada. Resulta complejo conducir una entrevista en la que predomina un proceso circular o en espiral, dentro del cual se van tocando una y otra vez los mismos temas con un cada vez mayor nivel de profundidad. El entrevistador poco experimentado puede encontrar grandes dificultades en la conducción de este proceso y puede suceder que el informante acabe decidiendo de qué temas habla y aportando escasa información sobre el problema que realmente motiva el diagnóstico.

Una forma de evitar estos imprevistos es preverlos con antelación, a fin de pensar las estrategias que pondremos en marcha. Para ello resulta de utilidad diseñar un esquema inicial de lo que esperamos sea la entrevista. En el anexo 4 hacemos una propuesta de cómo puede enfocarse la elaboración de este guión inicial. La mayoría de los autores (e.g. TAYLOR y BOGDAN, 1990; WOODS, 1989) consideran necesario planificar una *guía de la entrevista* para asegurarse del conjunto de temas que deben ser explorados. En todo caso, es evidente que no se trata de un protocolo estructurado que pretende garantizar una aplicación idéntica a distintos entrevistados potenciales. Será dentro de la situación concreta de cada entrevista cuando el entrevistador decida cómo explorar esos temas mediante preguntas y en qué momento.

Para la preparación de este tipo de entrevistas nos viene bien tener en cuenta la aportación de WU (1967[6], cit. por PEÑATE, 1993) que, aunque referida más bien a las entrevistas en general —y no a las entrevistas abiertas de forma particular—, nos resulta de gran utilidad para afrontar el proceso de diseño y preparación de este tipo de entrevistas. Este autor diferencia entre dos tipos de preguntas:

— *Preguntas-guías*, que hacen referencia a unas preguntas genéricas orientadas a conocer áreas concretas objeto de la exploración.

— *Preguntas de apoyo*, que son preguntas específicas sobre aspectos concretos que se hacen para aclarar o completar las preguntas-guía.

Así, una entrevista abierta encaminada a indagar sobre una cuestión puede partir de un conjunto de temas (o ejemplos de posibles preguntas-guías) y, a la luz de las respuestas dadas por el entrevistado, formular nuevas preguntas no previstas inicialmente que van a permitir que el entrevistado aclare lo que dice, ponga ejemplos, profundice, continúe en una determinada dirección...

Estas preguntas-guías no tienen por qué formularse según un orden concreto, sino de la forma más espontánea posible dentro del hilo de la conversación que está teniendo lugar. El llevarlas previstas de antemano permite que no se olviden temas importantes y que no se produzcan instantes de confusión en los que parezca que el entrevistador no sabe qué preguntar. Su uso adecuado dentro de la entrevista nuevamente dependerá de la pericia del entrevistador.

En todos los casos, las preguntas-guías deben estimular la libre expresión del informante y ser amplias en su formulación, a fin de no condicionar la respuesta en un sentido determinado. PATTON (1990) considera que la forma en que están formuladas las preguntas en las entrevistas abiertas constituye un tema crítico que condiciona la calidad de la información recogida. Así, en este tipo de entrevistas, las preguntas deben ser *verdaderamente abiertas*, lo cual quiere decir que han de permitir que el entrevistado se exprese en sus propios términos. Las preguntas *verdaderamente abiertas* no proponen una dimensión (sentimiento o ideas) en torno a la cual el entrevistado deba responder, sino que le permite seleccionar su respuesta de entre el conjunto de respuestas posibles en su repertorio. Precisamente esto es lo que le interesa al entrevistador: ver qué temas usa el entrevistado para explicar su realidad, conocer qué aspectos le resultan especialmente significativos y relevantes.

Un ejemplo quizá sirva para aclarar mejor cuándo una pregunta es *verdaderamente abierta*, usando la terminología propuesta por PATTON (1990). Si preguntamos a alguien «¿cómo de satisfecho está con el programa?», aunque inicialmente no ofrecemos alternativas que condicionen la respuesta, estamos dejando muy clara la dimensión en la que el informante debe responder: el grado de satisfacción respecto al programa. Ahora bien, es posible formular esta pregunta sin marcar una dirección desde la cual contestar: «¿cómo se siente ante el programa?, ¿cuál es su opinión sobre el mismo?». Esta nueva forma de preguntar permite que el entrevistado saque a la luz los aspectos más significativos de su experiencia con el programa y no exclusivamente su grado de satisfacción.

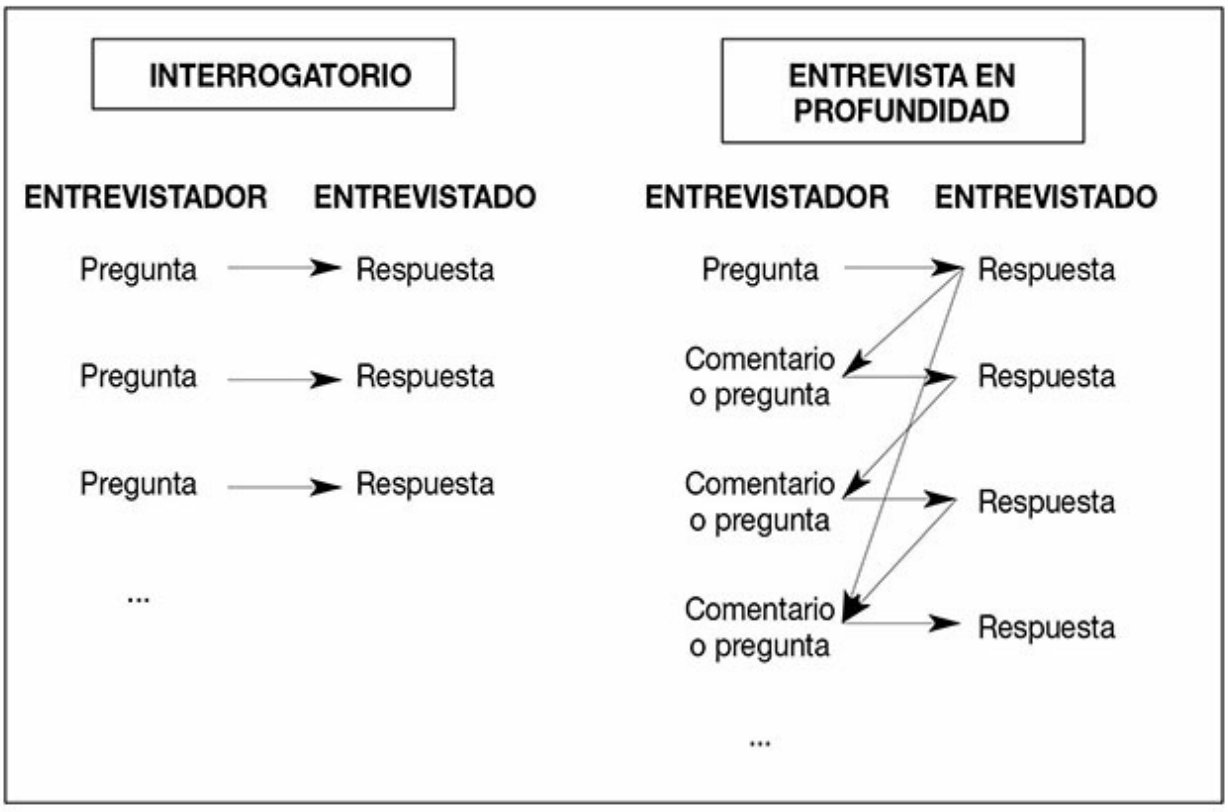
Por tanto, la guía que elaboremos para preparar la entrevista debe contener preguntas generativas, susceptibles de provocar respuestas profusas y peculiares. Las preguntas que pueden ser contestadas con un simple «sí» o «no» o con una breve frase no estimulan al informante a responder libre y detalladamente. Las preguntas indirectas, abiertas, invitan al informante a expresar ideas, a dar información que quizá puede no haberse conseguido

mediante acercamientos más directos. Si abusamos en nuestra guía de preguntas como «¿te gusta tu trabajo?», frente a «¿qué aspectos te gustan más de tu trabajo?», es posible que la entrevista acabe pareciendo un interrogatorio, o siguiendo un esquema muy similar al que conllevaría una entrevista estructurada.

En la figura 1 hemos intentado recoger esta idea. Frente a un proceso lineal donde cada pregunta conduce a una respuesta cuyo contenido no se refleja en la siguiente cuestión, se trata de promover una comunicación bi-direccional, en la que la formulación de las preguntas genera un proceso circular o en espiral que favorece la libre expresión del entrevistado. A partir de lo que el entrevistado responde, el entrevistador hará comentarios o nuevas preguntas que le permitirán profundizar más en cada cuestión, retomando cuando sea necesario las respuestas a preguntas anteriores.

FIGURA 1

ESQUEMA LINEAL Y COMUNICACIÓN BI-DIRECCIONAL EN EL DESARROLLO DE LA ENTREVISTA ABIERTA



A través de este esquema, la entrevista transcurre con fluidez y tanto informante como entrevistador se sienten más cómodos. Este último se muestra como un interlocutor animado que demuestra interés ante lo que dice el informante y hace comentarios a las respuestas de éste, comentarios que en ocasiones van convirtiéndose en nuevas preguntas (formuladas o no de forma interrogativa) que prolongan fluidamente la línea de

indagación. Por otra parte, las preguntas abiertas relajan la ansiedad en tanto que dan al informante libertad de acción a la hora de contestar. Las preguntas más cerradas o directas (que requieren una contestación rápida) tienden a potenciar la tensión dentro de la entrevista.

Otro aspecto que hay que cuidar en la preparación de la entrevista, cuando se están formulando las posibles preguntas-guías, se refiere al contenido de las preguntas. Si bien cualquier tema puede ser objeto de una cuestión formulada por el entrevistador, conviene que en la entrevista abierta no haya muchas preguntas sobre determinadas temáticas. Así, no es apropiado utilizar esta modalidad de entrevista para preguntar al entrevistado qué opina el colectivo o grupo a que pertenece. Para conseguir esta información, la entrevista estructurada o el cuestionario pueden resultar de mayor utilidad, sin provocar el efecto indeseado de que el entrevistado crea que le hacemos perder el tiempo por preguntar lo que piensan «otros» y no lo que piensa él mismo.

Así, un exceso de preguntas del tipo: «¿qué hacen los estudiantes de esta facultad cuando tienen que negociar un examen?», «¿cómo valoran los trabajadores de la empresa el clima que existe en ella?»... pueden conducir a que el entrevistado dé opiniones generales, sin profundizar mucho en su perspectiva personal del asunto, que es lo verdaderamente relevante en esta modalidad de entrevista. No se trata tanto de que nos diga qué piensan otros, sino qué piensa, hace, siente y opina él mismo. «¿Cómo es para tí un día de trabajo típico?, ¿podrías ponerme un ejemplo de situaciones en las que te sientes incómodo con tus colegas?»... son preguntas focalizadas en el informante y su experiencia y, además de aportar información peculiar de cada entrevistado, nos permiten obtener descripciones copiosas y detalladas, ya que animan a que la gente hable.

Por otra parte, en la preparación de la entrevista hay que tener en cuenta que conviene formular preguntas referidas a *cómo* y *por qué*, y no sólo preguntas de *qué*, *cuándo*, *dónde* o *quién*. Estas últimas nos van a informar de qué actividades practica y realiza el entrevistado y en qué condiciones. Sin embargo, no nos darán información sobre por qué las hace o cómo las hace.

En definitiva, la entrevista en profundidad debe constituirse en un proceso dinámico para lo cual es de gran ayuda preparar un conjunto de preguntas que faciliten la puesta en marcha y desarrollo de este proceso. Por lo tanto, las entrevistas abiertas o no estructuradas en realidad no lo son, en el pleno sentido de la palabra. Es decir, siempre hay un propósito que orienta su desarrollo. La existencia de este propósito explícito es el argumento que muchos autores (por ejemplo, SPRADLEY, 1979) utilizan para diferenciar la entrevista como técnica de diagnóstico o de investigación de la mera conversación entre dos o más personas. Por ello, siempre es útil tener preparada una guía que oriente la entrevista. Ahora bien, en su desarrollo, las entrevistas abiertas proporcionarán la información buscada «en su propio orden y en su propio tiempo» (WOODS, 1989:93).

2.1.2. Desarrollo y aplicación de la entrevista no estructurada

Dado que la entrevista abierta es un proceso dinámico, y en ella se van tocando los temas con cada vez mayor profundidad, resulta difícil sintetizar en pocas palabras cuáles son las reglas o los consejos que pueden favorecer la realización de la entrevista. A través de ella, el entrevistado comunica su experiencia y sus creencias, lo que a veces implica tocar temas o cuestiones de carácter íntimo y privado sobre los que no siempre se está dispuesto a hablar. Por eso, establecer un sentimiento de confianza en la relación de entrevista es un paso básico que permitirá que el entrevistado aporte información relevante (WOODS, 1989).

Esta confianza no surge de forma espontánea, sino que es la consecuencia de un proceso que tiene lugar a lo largo de la entrevista. Los primeros momentos de toda entrevista se caracterizan más bien por lo contrario: la desconfianza mutua entre entrevistador y entrevistado. Así, al principio de la entrevista es habitual que tanto uno como otro desconozcan cómo son vistos por su interlocutor, qué impresión causan, cómo se interpretan sus opiniones... Si bien este sentimiento de aprensión es natural al inicio de la entrevista, es conveniente superarlo si se quiere llegar a conseguir una información de calidad.

Comenzar la entrevista con una charla informal sobre temas en ocasiones intrascendentes ayudará a que el informante se relaje y adquiera seguridad en sí mismo. Al comienzo de la entrevista, es natural que se dé una cierta dosis de angustia que hay que resolver para que esta situación no incida negativamente en el desarrollo de la misma. Hacer que el entrevistado hable de temas generales de su experiencia cotidiana es la estrategia más útil para estos momentos iniciales. Por un lado, el entrevistado percibe que las preguntas que le hace el entrevistador se refieren a temas que conoce, a cuestiones en las que él es experto. Por otro, al hablar de su experiencia, la tensión se irá relajando y el entrevistador tendrá la oportunidad de mostrar su actitud de escucha para alentar progresivamente la confianza, repitiendo sus explicaciones o utilizando sus propios términos para formular preguntas, apoyando lo que dice sin emitir juicios de valor, denotando sensibilidad ante la expresión del informante, etc.

El establecimiento de una relación inicial adecuada pasa, según ALONSO (1994), por la puesta en práctica de una *empatía controlada*. Bajo este término, se refiere a un *ritual* (gestos, expresiones corporales, elección de determinadas palabras...) que tiende a estabilizar las tendencias disruptivas de la comunicación, favoreciendo la creación de un clima de naturalidad y neutralidad que permitirá la libre expresión de ideas y significados.

Por otra parte, TAYLOR y BOGDAN (1990) consideran que no es adecuado plantear al comienzo de la entrevista preguntas directivas. Ello crearía en los informantes una especie de idea respecto a aquello sobre lo que es importante hablar, abortando la posibilidad de tocar algunos temas. Por eso, al principio de la relación con los informantes es más adecuado «aparecer como alguien que no está totalmente seguro de

las preguntas que quiere hacer y que está dispuesto a aprender con los informantes» (p. 15).

La actitud de escucha es un elemento esencial en la marcha de la entrevista abierta. No sólo implica lo que WOODS (1989) denomina «aparentar que se escucha», es decir, exhibir un lenguaje corporal que denote la escucha atenta, sino que, en nuestra opinión, implica también focalizar la entrevista demostrando que se ha escuchado todo lo que ha dicho el informante. Por ello, hay que prestar atención para no cometer el craso error de formular preguntas en las que resulta evidente que el entrevistador no estaba escuchando el contenido del discurso del informante, además de utilizar el lenguaje del entrevistado, sus términos y expresiones, sus argumentos, los temas que toca... cada vez que se formula una nueva cuestión.

Aunque no existen recetas ni pautas simples que puedan asegurarnos el éxito de una entrevista abierta, algunos autores dan una serie de consejos que contribuyen a crear la atmósfera adecuada. Así, TAYLOR y BOGDAN (1990), resumen estos consejos en los siguientes puntos:

— *No emitir juicios de valor* ante lo que manifiesta el informante. Se trata de hacer ver que se le acepta y comprende, pero no que se le está juzgando.

— *Permitir que la gente hable* es importante que el informante perciba la libertad con la que puede expresarse delante del entrevistador.

— *Prestar atención*, comunicando un interés sincero en lo que el informante está diciendo.

— *Ser sensible*. Las palabras y gestos del entrevistador deben hacer gala de sensibilidad ante el informante. Esto implica saber ser simpáticos sin llegar a la condescendencia, mantenerse alejado de heridas abiertas, «hacerse el tonto» sin llegar a resultar insultante, etc.

Permitir la libre expresión, no obstante, no es tan fácil como parece. En ocasiones, origina que el entrevistado se extienda en temas que no son de gran interés. En los primeros momentos de la entrevista, no conviene interrumpir al informante. Por un lado, este tipo de divagaciones resultan normales en un tipo de entrevistas abiertas en las que no se ha partido de unas pautas y reglas organizativas fijas. Por otro, como ya hemos planteado, mantener hablando al informante contribuye a generar el clima de confianza necesario para la buena marcha de la entrevista. Ahora bien, en momentos posteriores de la entrevista, sí resulta conveniente reconducir la conversación. Gestos sutiles como dejar de asentir y de tomar notas pueden ayudar, así como cambiar suavemente el tema tras una pausa en la conversación. Como plantean TAYLOR y BODGAN (1990), con el tiempo el informante aprende a leer nuestros gestos y está al tanto, cada vez más, de aquello que nos interesa, por lo que él mismo percibe cuándo debe cerrar un tema.

En general, la dinámica de la entrevista se orienta hacia la descripción en profundidad de los detalles y las experiencias de las personas, así como de los significados que éstas

les atribuyen. Estos significados son relativos y varían de una persona a otra, por lo que no es aconsejable dar las cosas «por sentadas», sino como plantean TAYLOR y BOGDAN (1990), aún a riesgo de parecer ingenuo, el entrevistador debe pedir constantemente al informante que clarifique y elabore lo que ha dicho. SPRADLEY (1979) también se refiere a esto afirmando que el entrevistador tiene que enseñar a los informantes a que se comporten como tales.

Pasando ya a analizar cómo deben plantearse las intervenciones del entrevistador, conviene recordar que la situación de entrevista ha de superar el esquema de un interrogatorio. El que no se disponga de un guión organizado y estructurado de preguntas, no significa que no exista un marco que focalice la entrevista. Si así fuera, la entrevista se convertiría en una *ensalada de palabras* (ALONSO, 1994). Por ello, las intervenciones del entrevistador deben contribuir a crear el clima y la relación adecuados. La siguiente cita de ALONSO (1994:233-234) expresa con un alto grado de claridad esta idea:

> «Se trata de que durante la entrevista la persona entrevistada produzca información sobre todos los temas que nos interesan, pero no de ir inquiriendo sobre cada uno de los temas en un orden prefijado. El objetivo es crear una relación dinámica en que, por su propia lógica comunicativa, se vayan generando los temas de acuerdo con el tipo de sujeto que entrevistamos, arbitrando un primer estímulo verbal de apertura que verosímilmente sea el comienzo de esa dinámica que prevemos.»

Saber dinamizar la situación de entrevista implica elegir el momento adecuado para cada tipo de intervención. ALONSO (1994) diferencia entre seis tipos básicos de intervenciones del entrevistador:

— *Complementaciones*. Intervenciones que estimulan un discurso narrativo y descriptivo, tratando de abundar en su exhaustividad y profundidad.

— *Interpretaciones*. Intervenciones que pretenden expresar una actitud del entrevistado no explicitada.

— *Preguntas sobre el contenido*. Intervenciones interrogativas que solicitan una identificación suplementaria.

— *Preguntas sobre la actitud*. Intervenciones también interrogativas que solicitan la identificación de las actitudes.

— *Eco*. Intervenciones que aíslan reiterando una parte a fin de subrayar su importancia.

— *Reflejo*. Es la reiteración que refleja en el entrevistado la actitud del que habla. A través de él se persigue que el entrevistado se refiera de forma más amplia a su posición personal y se centre en su propio pensamiento.

Como se sugirió en el apartado anterior, los comentarios del entrevistador pueden cumplir la misma función que las preguntas. Así, en lugar de preguntar «¿cuáles son los problemas más difíciles a los que te enfrentas en tu trabajo?», se puede hacer un comentario abierto del tipo: «Me imagino que en un trabajo como el tuyo es habitual que se presenten problemas difíciles de resolver...». Este comentario suscitará la respuesta

del informante, cumpliendo el mismo fin que la pregunta, pero aportando mayor dinamismo y naturalidad a la situación de entrevista. Algunas veces, para conseguir una mayor profundización en la respuesta, se puede utilizar la ya mencionada estrategia del eco. Así, ante expresiones del informante que consideramos de interés para profundizar, repetimos un segmento textual de lo que ha dicho en tono interrogativo, indicando de esta forma que nos interesa que profundice en ese aspecto. Por ejemplo, si el entrevistado dice «... estoy satisfecho de la relación con mis colegas, pero no soporto a los "fantásticos" de turno», una forma de ayudarle a que matice el significado es repetir con tono interrogativo la expresión que nos ha causado sorpresa («¿los "fantásticos de turno"?»). Tiene el mismo efecto hacer una pausa en la conversación, sin decir nada, manteniendo una expresión interrogante en la mirada.

Los silencios son también un elemento importante en las entrevistas no estructuradas. A través del uso de los silencios se pueden conseguir distintos fines: invitar al entrevistado a que siga hablando, sin interferir en lo que está diciendo; provocar o incitar su participación, haciendo que se sienta incómodo con el silencio y estimulándolo así a hablar; etc. El silencio puede servir también justo para lo contrario: anular o cortar intervenciones que se van alejando del centro de interés, dejando de reforzar la conducta del entrevistado. Respecto a los silencios, RODRÍGUEZ SUTIL (1994) comenta que suelen provocar ansiedad en el entrevistado, pero eso no tiene por qué ser en sí mismo algo negativo, sino que cierto nivel de ansiedad puede ayudar a que su estilo básico aflore de manera más clara

Además de estas intervenciones por parte del entrevistador, en general, son necesarios lo que TAYLOR y BOGDAN (1990:127) denominan como «controles cruzados». Esto consiste en hacer comprobaciones de las historias y narraciones de los informantes, examinando la coherencia de lo dicho por éste en diferentes momentos. En ocasiones esto puede implicar comprobar los datos con otras fuentes, bien sean documentales (registros oficiales, u otro tipo de documentos) o personales (otras personas que corroboran informaciones). Aparte de las implicaciones que este tipo de comprobaciones cruzadas pueden tener para la fiabilidad y validez de la entrevista, resultan útiles para aumentar la cantidad y la calidad de la información recogida durante el desarrollo de la misma. Así, volver una y otra vez a un tema, haciendo ver al informante que no nos han pasado por alto los aspectos «oscuros» o las posibles incoherencias es lo que permite que seamos capaces de conseguir una información completa, profunda y exhaustiva. Nuevamente nos vuelve a aparecer la idea de circularidad o de proceso en espiral que acompaña a la entrevista abierta.

Respecto a la finalización de la entrevista, no se recomienda ni la terminación brusca y precipitada, sin haberse alcanzado la información deseada y la profundidad adecuada de la misma, ni tampoco un «abuso» del tema y del entrevistado. No hay que olvidar que en momentos posteriores puede resultar necesario volver a solicitar alguna información. El entrevistado estará más o menos dispuesto según el recuerdo —agradable o desagradable— que le haya quedado de la entrevista.

No hay recetas para ser un buen profesional de las entrevistas abiertas. Esto depende de muchos factores: la propia habilidad, sensibilidad, experiencia... La personalidad y habilidades del entrevistador (curiosidad, espontaneidad...) constituyen un factor de gran importancia para determinar la calidad de la información obtenida a través de la entrevista abierta (MERRIAM, 1988). Lo que sí parece claro es que la capacidad crítica y de reflexión del entrevistador sobre su propia experiencia es la llave que puede ponerle en el camino adecuado. Como dicen COHEN y MANION (1990), el entrevistador debe desarrollar la capacidad para evaluar continuamente la entrevista mientras está en marcha, a fin de obtener datos significativos. Y es que la figura del entrevistador es central en este tipo de entrevistas. Como plantean TAYLOR y BOGDAN (1990), lejos de asemejarse a un robot recolector de datos, en la entrevista en profundidad el propio entrevistador es el instrumento de la investigación, y no lo es un protocolo o formulario de entrevista. Su rol no consiste sólo en obtener respuestas, sino también en aprender qué preguntas hacer y cómo.

2.1.3. El registro de la información en la entrevista abierta

Es evidente que en la entrevista abierta se genera mucha información que resulta difícil de registrar en forma de notas. Esto, por otra parte, llevaría a que el entrevistador estuviera demasiado centrado en su cuaderno de notas, lo que dificultaría la creación del clima y la dinámica adecuados en la entrevista. Por ello resulta necesario utilizar la grabación de las producciones verbales del entrevistado. Como se planteó en el capítulo dedicado a la observación, la grabadora —o los medios tecnológicos de registro de información, en general— pueden crear ciertos recelos en la persona entrevistada. No obstante, lo habitual es que nadie se niegue a que se use. Si una persona ya ha accedido a realizar la entrevista, es consciente de la importancia que la grabación va a tener para la recogida de datos del diagnóstico, por lo que es inusual encontrar reluctancias a su uso. Por otra parte, estas entrevistas se prolongan considerablemente en el tiempo, por lo que, aunque al principio, puede estar cohibido por la grabación, posteriormente llegará a centrarse en la conversación y olvidar la presencia de la grabadora.

En cualquier caso, antes de usarla es conveniente aclarar que, en cualquier momento, el entrevistado tendrá en sus manos la posibilidad de parar la grabación, en caso de que no quiera que ciertas informaciones queden grabadas. Asegurar el anonimato del entrevistado y la confidencialidad de la información que aporta son pasos necesarios para obtener el permiso de la grabación.

Por otra parte, existen un conjunto de reglas, más o menos obvias, que conviene tener presentes cuando se usa una grabadora. La primera de ellas ya fue expuesta en el capítulo anterior y se refiere al uso de equipos poco aparatosos que tengan un impacto mínimo en la situación de entrevista. Lo ideal es que, tras pocos minutos de entrevista, el entrevistado se olvide de la existencia de la grabadora; en la medida en que ésta sea grande o contenga elementos invasivos del espacio personal (por ejemplo, un micrófono

junto a la boca del entrevistado), seguirá siendo un elemento extraño en la entrevista, cuya presencia resulta difícil de olvidar.

Otras cuestiones más obvias, pero no menos importantes, se refieren a comprobar el estado del equipo antes de la grabación. Se trata de ver si tenemos pilas y cintas suficientes, además de revisar que el equipo funciona correctamente. Es conveniente también elegir un lugar tranquilo en el que no haya interrupciones ni ruidos de fondo que dificulten la grabación. Lo peor que podría sucedernos es desarrollar una entrevista y comprobar, tras su finalización, que no se ha registrado nada. Esta situación ya no tiene marcha atrás, puesto que en este caso resulta poco adecuado y poco profesional repetir la entrevista.

Una vez realizada y grabada la entrevista, ha de procederse a la transcripción de la misma. Es importante reflejar con fidelidad el lenguaje del informante, sus entonaciones, pausas, titubeos... La persona que ha realizado la entrevista es la más adecuada para transcribirla, ya que es la que mejor puede recrear el ambiente de la misma. Si ella no la transcribe, al menos debe escuchar la cinta y comprobar la fidelidad de la transcripción.

Por otra parte, en la transcripción de cada entrevista se debe indicar la fecha y hora en que se realizó, identificando quién habla en cada momento (el entrevistador o el entrevistado), y dejando amplios márgenes que faciliten el posterior análisis de la información.

A pesar de que se realice un registro magnetofónico, no debe desestimarse completamente el uso de notas. Si bien el discurso oral está bien recogido a través de la grabadora, los gestos, las reacciones y los matices de la expresión no lo están y pueden ser de gran ayuda para comprender lo que el entrevistado piensa. Para algunos autores (FONTANA y FREY, 1988), los elementos no verbales son tan importantes en una entrevista como los verbales. Por ello, conviene disponer de una libreta que debe ser colocada estratégicamente a fin de no interrumpir el ángulo visual entre entrevistador e informante. En ella se pueden apuntar palabras clave o frases sin seguir un orden temporal, para que el informante no esté demasiado pendiente de qué es lo que está registrando el entrevistador. El uso de estas notas también será de gran ayuda para evocar qué temas se han tratado, cuándo y cómo, si resulta necesario volver atrás y retomar temáticas o aspectos de interés.

La toma de notas de forma discreta no tiene por qué disminuir la naturalidad o espontaneidad de la entrevista y puede resultar de gran ayuda para evocar el recuerdo de determinados aspectos o situaciones de la conversación.

2.2. Entrevistas estructuradas

De forma general se pueden definir como entrevistas que «no dejan nada al azar» (PEÑATE, 1993:184). En ellas, el entrevistador dispone de un protocolo en el que se

recogen todas las preguntas a realizar así como, en algunos casos, las posibles respuestas a elegir por parte del entrevistado. Por esta razón, es habitual definir este tipo de entrevista como «cuestionario realizado oralmente en presencia del entrevistador» (BUENDÍA, 1994:207).

Este tipo de entrevistas se utiliza para obtener información de una muestra relativamente amplia. Aunque dependerá en gran medida de los objetivos del diagnóstico, si el número de sujetos que intervienen en el mismo es muy limitado (apenas cinco o diez personas), es preferible usar una modalidad de entrevista más abierta susceptible de aportar una información más profunda y detallada. Sin embargo, si el número de sujetos implicados es mayor, conviene utilizar entrevistas estructuradas o semi-estructuradas que resultarán más fáciles de aplicar y de analizar.

La entrevista estructurada resulta útil para analizar situaciones desde una óptica exploratoria y sondear las opiniones de un colectivo ante las mismas. Al ser una técnica estructurada exige que se delimite *a priori* el marco conceptual desde el que se realiza la entrevista. Sólo desde este marco conceptual tendrán sentido las preguntas y alternativas de respuesta, las cuales reflejan la reflexión y opinión del entrevistador ante el problema o situación que pretende evaluar. Por lo tanto, se trata de una técnica de recogida de información que se utiliza para comprobar los supuestos previos mantenidos por la persona que realiza el diagnóstico, y no tanto para conocer los esquemas de otras personas (como sucede en las entrevistas abiertas). La respuesta de los sujetos a las preguntas de la entrevista estructurada, en el fondo, permiten reforzar o desestimar los supuestos previos de los que parte el entrevistador.

Por todo lo anterior, puede deducirse que el proceso de construcción de las entrevistas estructuradas es muy similar al expuesto en el capítulo precedente en relación con los sistemas categoriales para la observación. Dadas las similitudes entre las entrevistas estructuradas y los cuestionarios, el proceso de construcción de ambos instrumentos se expondrá conjuntamente en el apartado dedicado a los cuestionarios como procedimientos diagnósticos.

2.2.1. Diferencias entre la entrevista estructurada y el cuestionario

A fin de caracterizar mejor las entrevistas estructuradas, y puesto que hemos definido esta técnica como «cuestionario oralmente administrado», consideramos necesario analizar las diferencias existentes entre las entrevistas estructuradas y los cuestionarios, lo que nos permitirá conocer más a fondo las técnicas en cuestión.

Aunque son muchas las semejanzas que comparten el cuestionario y la entrevista estructurada, existen diferencias entre ambas técnicas que conviene resaltar. Estas diferencias provienen básicamente de lo que «diferencia» —valga, por el momento, esta redundancia— a ambos procedimientos. Así, en el *cuestionario*, la información se ha obtenido sin que medie la relación personal entre sujeto y encuestador; el primero contribuye con sus respuestas escritas a las preguntas formuladas por el segundo. En

cambio, en la *entrevista*, al estar presente el entrevistador en la formulación de las instrucciones y preguntas —pese a que hayan sido preparadas con antelación—, existe una mayor flexibilidad en conseguir la información (si tiene alguna duda, el entrevistado puede formularla antes de contestar). Asimismo, el entrevistador tiene la posibilidad de observar al sujeto y la forma en que se enfrenta a la situación de entrevista.

El cuestionario es un procedimiento menos costoso que la entrevista. Los cuestionarios pueden ser administrados a un gran número de sujetos simultáneamente, mientras que la entrevista requiere, por lo general, un interrogatorio individual y por separado de cada sujeto. Con cierto humor, SELLTIZ, et al. (1976:270) resumen esta misma idea al afirmar que «los cuestionarios pueden ser enviados por correo; los entrevistadores, no»... Otra ventaja atribuible a los cuestionarios respecto de las entrevistas es que en los primeros, la sensación de anonimato es mayor. El sujeto puede sentirse más libre para expresar sus opiniones. En la entrevista, pese a que se asegure el uso confidencial de toda información aportada, el sujeto tiene que verbalizar su respuesta frente al entrevistador. En ellas, además, es posible que el entrevistador conozca el nombre y dirección del entrevistado, lo que puede plantear dudas —por irracionales que sean— respecto a la posibilidad de identificación personal de sus opiniones.

SELLTIZ et al. (1976) resaltan otra ventaja del cuestionario frente a la entrevista. Mientras que en la segunda, el sujeto se ve en la necesidad de dar una respuesta inmediata —ante el entrevistador que tiene en frente—, en el cuestionario, el sujeto dispone de un margen amplio de tiempo para considerar cada punto cuidadosamente antes de contestar con la primera impresión. Claro que esto, en nuestra opinión, puede resultar un arma de doble filo: por un lado, porque, esperando a tener tiempo, el cuestionario puede quedar sepultado bajo una montaña de papeles y no ser reenviado al encuestador; por otro, si el sujeto tiene más tiempo para pensar y no contesta tan espontáneamente, los sesgos que afectan a estas técnicas de recogida de información pueden incidir con más facilidad en los cuestionarios que en las entrevistas.

Pese a las ventajas que el cuestionario puede presentar respecto a la entrevista, existen un conjunto de limitaciones del mismo que hacen aconsejable el uso de las entrevistas en ciertos casos. Esto es evidente, por ejemplo, cuando queremos encuestar a un determinado sujeto o grupo de sujetos que no saben leer y escribir o que tienen gran dificultad para expresarse por escrito. Por otra parte, ante la presencia del entrevistador es más probable que la persona colabore aportando la información que se solicita. El cuestionario, al ser administrado por correo, muchas veces acaba en la papelera, bien porque la persona no desea contestar —y no se siente «moralmente» obligada al no tener a nadie a quien dar excusas— o bien porque, aun deseando colaborar, no encuentra el momento justo para hacerlo y acaba olvidándose del instrumento que se le ha enviado. Así, siempre que en la realización de un diagnóstico empleemos un cuestionario, hemos de saber que los porcentajes de devolución escasamente superan el 80% del total de los cuestionarios enviados o distribuidos, siendo más bajo este porcentaje en la mayoría de

las ocasiones. No obstante, profundizaremos en esta cuestión en el apartado dedicado al cuestionario.

Otra ventaja importante a destacar de la entrevista y que ya hemos apuntado, es su mayor flexibilidad. En ella, si el sujeto no ha entendido algo o denota que interpreta incorrectamente la pregunta, es posible hacer una reformulación de su enunciado a fin de facilitar su comprensión (aunque el lector encontrará en apartados posteriores de este capítulo algunas matizaciones sobre cómo proceder en estos casos). Esta peculiaridad de la entrevista la convierte en un medio muy útil para poner a prueba los cuestionarios. Así, en los estudios piloto que se realizan sobre la versión inicial de un cuestionario, resulta de gran utilidad administrarlos bajo el formato de entrevista. De esta forma, el profesional del diagnóstico puede percibir de primera mano las dificultades que plantean las preguntas, las reacciones que provocan... pudiendo de esta forma, mejorar el instrumento para futuras aplicaciones.

2.2.2. Preparación de la cédula

Como ya se ha indicado, la fase de diseño y preparación de la entrevista estructurada es la que tiene mayor importancia. En ella, además de definirse escrupulosamente lo que se va a evaluar, se preparan todas las preguntas y se prevén las dificultades que pueden presentar cada una de ellas, así como las posibles incidencias que pueden tener lugar durante el desarrollo de la entrevista. El objetivo a alcanzar es que las condiciones de aplicación de la entrevista a cada sujeto sean idénticas, de tal forma que se minimicen los efectos debidos al entrevistador y la situación de entrevista. Si la entrevista no es aplicada en igualdad de condiciones a todos los sujetos, se pueden introducir sesgos que afectan a la calidad de la información recogida.

Dado que todo está rigurosamente controlado y planificado tras la fase de diseño, en las entrevistas estructuradas bastará con un mínimo entrenamiento de los entrevistadores, sin ser necesario que éstos pongan en juego las complejas habilidades que se requieren en las entrevistas abiertas y en profundidad.

Puesto que, a fin de no repetir las mismas ideas, el proceso de diseño y construcción de las entrevistas abiertas se expone en el apartado dedicado al cuestionario, nos vamos a limitar aquí a describir los aspectos que hay que tener en cuenta a la hora de preparar la cédula. El lector debe entender que dicha cédula no surge de la nada, sino que es el producto de las tareas y decisiones relativas al diseño que se verán más adelante. Sólo cuando estas tareas han sido realizadas tiene sentido disponerse a preparar la cédula o formato escrito de la entrevista.

Por tanto, cuando hablamos de la cédula de una entrevista nos referimos al esquema de la misma operativizado en una hoja de papel que contiene todo lo necesario para aplicar la entrevista. La cédula, que debe ser cumplimentada por el propio entrevistador y no por el sujeto, es el documento en el que aparecen todas las instrucciones a dar, las preguntas a formular y sus opciones de respuesta, así como los espacios en blanco o

casillas vacías para consignar la respuesta del sujeto. Varias son las funciones que cumple la cédula. HAGUE y JACKSON (1994) destacan las más importantes:

— Sirve de guía en la entrevista en tanto que indica las preguntas a realizar y el orden de las mismas.

— Sirve como garantía de que a todas las personas entrevistadas se les pregunta lo mismo y de la misma forma.

— Sirve como formulario para anotar y recoger las respuestas.

Al igual que el cuestionario, la entrevista estructurada debe comenzar con una breve introducción en la que se hace referencia a los fines de la misma y el proceso a desarrollar. Esta introducción, así como las instrucciones necesarias para realizar la entrevista (por ejemplo, relativas a cómo debe responder el sujeto según el tipo de preguntas), deben figurar en la cédula, puesto que han de exponerse de forma estrictamente igual a todos los sujetos a entrevistar con el fin de evitar posibles sesgos en la información.

La presentación de la entrevista (introducción e instrucciones) debe ser breve, pues presentaciones extensas pueden resultar, además de aburridas, sospechosas para el entrevistado. En ella se debe mencionar la información relevante sin usar expresiones o construcciones gramaticales rebuscadas y complejas. Esta información relevante se refiere a: los objetivos que se persiguen con la entrevista, los fines más amplios del proceso de diagnóstico en el que la entrevista se inserta y las instrucciones para responder a las preguntas. Además, es necesario asegurar la confidencialidad de la información aportada por el entrevistado.

Tras la presentación, figurarán las preguntas a formular y las alternativas de respuesta en cada una de ellas, con los espacios o casillas oportunos para registrar la respuesta aportada por el sujeto. Al final de la cédula, se debe consignar la fórmula de agradecimiento.

En la preparación de la cédula, resulta de especial utilidad diferenciar con tipos de letras distintos las diferentes instrucciones, preguntas, contenidos, etc. (MANZANO, ROJAS y FERNÁNDEZ, 1996). Así, conviene diferenciar entre la información que va a ser leída ante la persona entrevistada y la que está reservada para el encuestador. El uso de diferentes tipos de letras, colores y métodos de resaltes (negrilla, cursiva...) servirá a tal fin.

Esta estrategia de diferenciar visualmente la información tiene mucha utilidad para el ulterior desarrollo de la entrevista. En las entrevistas estructuradas el entrevistador no debe improvisar cuando el sujeto manifiesta no haber entendido una pregunta o demuestra escaso interés por responder. Si el entrevistado improvisa una nueva forma de formular la pregunta, está cambiando las condiciones de la entrevista para ese sujeto. Puesto que las preguntas deben ser las mismas para todos los sujetos, esto conlleva un importante sesgo en la entrevista, máxime si tenemos en cuenta que, en nuestro afán por

conseguir la respuesta, podemos formular el enunciado de la pregunta de tal forma que induzcamos a que nos contesten en una dirección determinada. Por ello, resulta adecuado prever esta posibilidad (que el sujeto no quiera o no sepa responder) y llevar preparada una información alternativa para estas situaciones, información que siempre será la misma cuando un sujeto manifiesta su falta de comprensión del enunciado inicial. Al igual que en los casos anteriores, este nuevo tipo de información también debe ser resaltada con un tipo, color o resalte de letra distinto a los anteriores.

Elaborar la cédula bajo estos criterios permitirá que los entrevistadores estén preparados ante las posibles contingencias que puedan tener lugar durante la aplicación de la entrevista, sin que tengan que acudir a la improvisación.

2.2.3. Aplicación de la entrevista estructurada y el registro de la información

Algunos autores (por ejemplo, SILVA, 1992) diferencian dos momentos diferentes en las entrevistas: por un lado, el comienzo y, por otro, el cuerpo de la misma. Aunque en este tipo de entrevistas no se busca establecer una relación de confianza, ello no implica que la situación en que se desarrolla tenga que ser necesariamente fría y distante.

En general, el primer papel del entrevistador consiste en promover un buen inicio de la entrevista. Un saludo afectuoso, un apretón de manos firme y una presentación adecuada de sí mismo van a ayudar a establecer un buen *rapport* (MORGAN y COGGER, 1989). Estos gestos de cortesía inicial van a contribuir a que el entrevistado piense que la entrevista va a ser una experiencia agradable y no penosa.

Aunque es en la preparación de la cédula el momento en que se decide la forma de presentación, cuando se va a aplicar la entrevista, el encuestador debe, en primer lugar, presentarse, justificar por qué ésta es necesaria, dar a conocer sus finalidades e indicar quién realiza el diagnóstico o investigación y con qué fines. Esto no sólo permitirá suscitar una buena acogida despertando el interés, sino que también contribuirá a que no se levanten recelos o desconfianzas respecto al qué, al quién y al por qué del diagnóstico que se va a realizar.

Todo lo que el entrevistador dirá o preguntará al entrevistado está recogido en la cédula. Del mismo modo, todo lo que el entrevistado diga quedará registrado en la misma. A pesar de ello, el entrevistador experimentado sabe utilizar la cédula de manera informal, sin dar la impresión de un interrogatorio forzado. Formular las preguntas en un tono de voz natural, combinar rápidos vistazos a la cédula con miradas al entrevistado, utilizar frases de transición entre una pregunta y la siguiente («veamos ahora...»)... son algunas habilidades que facilitarán la realización de la entrevista. De todas formas, no se debe olvidar que las intervenciones del entrevistador están rigurosamente especificadas en la cédula, limitándose éste a leer lo que pone en ella. El efecto a conseguir es que la atención del entrevistador esté centrada en la persona entrevistada y no en la cédula que tiene delante (GARCÍA JIMÉNEZ, 1997).

Un aspecto de gran importancia en este tipo de entrevista se refiere al riesgo de sugerir o inducir la respuesta. Tal y como se ha visto con anterioridad, las entrevistas estructuradas han de aplicarse en situaciones análogas a todos los sujetos. En el caso en que el contenido de las preguntas se modifique en una o varias entrevistas, no podemos estar seguros de que todos los encuestados han partido de las mismas condiciones a la hora de responder. Sin embargo, en ocasiones, cuando el entrevistado no ha entendido bien la pregunta o no sabe qué contestar, o contesta un «no sé» con objeto de evitar pensar sobre el tema, es necesaria la intervención del entrevistador. En estas situaciones en las que es preciso reformular la pregunta inicial para obtener la respuesta, se corre el riesgo de cambiar su enunciado o reformularla de tal forma que induzca a responder en una dirección determinada. Por ello, una cédula en la que se ha previsto esta contingencia (con formulaciones alternativas resaltadas visualmente) resulta de gran ayuda para el momento concreto en que el entrevistador se enfrenta a la situación.

Para algunos entrevistados recalcitrantes puede resultar necesario comentarios o estrategias adicionales. Además de la formulación alternativa prevista en la cédula, existen otras técnicas que nos permiten indagar la respuesta, cuando ésta es incompleta o insuficiente (MANZANO, ROJAS y FERNÁNDEZ, 1996):

— *Silencios*. La situación de silencio es comprometida o incómoda, por lo que las personas tienden a «rellenar» los silencios dando más detalles acerca de su respuesta.

— *Eco*. Al repetir la respuesta —o parte de ella— de la persona entrevistada, ésta tenderá a asentir si es correcta y a completar su respuesta con más detalles para no provocar otro eco.

— *Asentimientos*, mediante movimientos de cabeza o monosílabos que animan a la persona a que continúe con la elaboración de su respuesta.

— Volver a *centrar la respuesta* con expresiones del tipo «¿cómo podríamos resumir su respuesta...?» o «concretando, en resumen, Vd. opina...».

— *Preguntas no directivas* como «creo que no le he entendido muy bien...», o «¿qué quiere decir con eso?».

— *Aclaraciones* sobre respuestas ya realizadas, ante las que se pide confirmación o matización de algún aspecto.

En entrevistas de este tipo que se aplican a diagnósticos individuales, existen otros procedimientos para completar la información (SILVA, 1992): el «recurso al ejemplo» real y reciente puede ayudar en ciertas entrevistas a aclarar y precisar algunos datos; las «estrategias de reconocimiento», que facilitan al sujeto información para que la reconozca y así facilitar el recuerdo, evitando sugerir las respuestas; y, reservar, antes de la terminación de la entrevista, unos momentos para discutir abiertamente con el sujeto las inconsistencias o vacíos de la información recogida.

En cualquier caso, no se debe olvidar que si el sujeto no quiere o no sabe responder mejor a una o varias preguntas de la entrevista, en un momento posterior se puede desestimar su participación en el estudio, bien de forma global en todas las respuestas que ha dado, bien de forma parcial, desestimando las respuestas que se prestan a la duda. Si esto sucediera con muchos sujetos, el problema será más grave, puesto que nos está indicando que la pregunta en cuestión está mal formulada o presenta dificultades para su correcta comprensión.

Respecto al *registro de la información*, se plantea un cierto dilema. Si se hace simultáneamente a la formulación de preguntas, aparentemente se dificulta el curso natural de la conversación, mientras que el registro posterior, conlleva una fuerte reducción y distorsión de la información. Autores como GARCÍA MARCOS (1983) recomiendan el uso de una estrategia mixta, de forma que se recojan datos importantes o ciertas expresiones del sujeto de forma textual durante la entrevista, y se complete la información recopilada inmediatamente después de su realización. Así, durante la entrevista, el entrevistador presta atención a la relación que establece con el entrevistado.

La tarea del registro de las respuestas será más fácil en tanto mayor sea el grado de estructuración de las preguntas. En los casos en que el sujeto tiene que elegir una opción de respuesta entre las alternativas ofrecidas, el registro consistirá en marcar la opción seleccionada, mientras que en las preguntas abiertas, la información a consignar es mayor. Incluso en entrevistas cuyas preguntas son cerradas, al terminar, conviene repasar que se han codificado las respuestas correctamente o si hay alguna omisión o despiste por parte del entrevistador. En respuestas abiertas, conviene utilizar frases literales. No se deben resumir las respuestas con las propias palabras o «pulirlas» para que parezcan más elaboradas.

3. EL CUESTIONARIO EN LOS PROCESOS DIAGNÓSTICOS

3.1. Concepto y características

Cuando el número de sujetos que pueden proporcionar información para el diagnóstico es numeroso, el cuestionario auto-administrado resulta una estrategia más económica que la entrevista estructurada. Como ya se ha planteado, las ventajas y desventajas del uso del cuestionario frente a la entrevista provienen del hecho de ser aquél una técnica basada en la interacción impersonal. Desde esta perspectiva, el cuestionario nos proporciona dos ventajas: por un lado, la posibilidad de acceder a más sujetos que si tuviéramos que entrevistarlos uno a uno, y por otro, la mayor normalización y estandarización de la situación, ya que por mucho cuidado que se ponga a esta cuestión en las entrevistas, siempre es más fácil que la situación de entrevista varíe de un entrevistador a otro y de un entrevistado a otro.

Sin embargo, estas ventajas derivadas de la relación impersonal entre encuestador y encuestado tienen unas contrapartidas que suponen limitaciones importantes del cuestionario a tener en cuenta. En primer lugar hemos de referirnos a la necesidad de formular las preguntas por adelantado, de forma que su intención y significado queden claros sin necesidad de explicaciones complementarias. El profesional del diagnóstico debe disponer los medios necesarios para que las preguntas no generen dudas o interpretaciones erróneas, ya que él no estará presente durante la administración del instrumento que diseña. Por esta razón, en el cuestionario es aún mayor la importancia del estudio piloto respecto a las entrevistas estructuradas.

Según MUCCHIELLI (1974), el porcentaje de respuestas fallidas (cuestionarios no devueltos) puede oscilar entre el 25 y el 80%. Por su parte, FOX (1981) asegura que el número de devoluciones rara vez supera el 50% y, en ocasiones, no sobrepasa el 30%. Por ello, su preparación (tanto lo referido a la preparación de las preguntas como a la redacción de la introducción) ha de ser aún más rigurosa que en el caso de las entrevistas estructuradas. En ellos, no va a estar presente un entrevistador que aclare el matiz de una pregunta o que indique cómo tiene que contestar el sujeto en caso de duda. Por otra parte, la mera presencia física del entrevistador contribuye a que la persona se sienta más comprometida a participar. En cambio, cuando el cuestionario llega por correo (o se ha repartido previamente), el sujeto no se siente tan obligado, y puesto que puede cumplimentarlo en cualquier momento, es posible que acabe olvidándose del mismo por otras ocupaciones.

Además de poner especial cuidado en su elaboración, es necesario realizar algunas operaciones que contribuirán a que el número de cuestionarios devueltos sea el mayor posible. Respecto a estas operaciones, FOX (1981) propone las siguientes:

1) Limitación de la extensión del cuestionario para que los sujetos tengan que dedicar el menor tiempo posible a contestarlo.

2) Estructuración del modelo de respuestas en el mayor grado posible para reducir al mínimo lo que tengan que escribir los sujetos.

3) Redacción del material introductorio de un modo elocuente y sincero para que los sujetos conozcan la finalidad de la investigación y el uso que se hará de los datos y se convenzan de que es una finalidad útil y profesionalmente deseable.

4) Arbitrar algún sistema para que los sujetos, si lo desean, puedan conocer los resultados del diagnóstico.

SELLITZ et al. (1976) también mencionan un conjunto de factores que influyen en el porcentaje de devolución de cuestionarios: la forma de presentación, la longitud, la naturaleza de la carta que lo acompaña, la facilidad para rellenarlo y devolverlo, etc. En la medida en que un cuestionario se presente de forma atractiva, breve y se perciba como fácil de cumplimentar y de devolver, el destinatario del mismo se verá más o menos motivado a colaborar. Aun controlando estos factores, un determinado número de cuestionarios no será devuelto. Este hecho conviene que sea tenido en cuenta por parte del profesional del diagnóstico, ya que según SELLTIZ et al. (1976), las personas que lo cumplimentan y devuelven suelen ser las más sedentarias (y, por lo tanto, reciben el cuestionario), las más interesadas, cultas y razonables.

Es necesario plantear aquí, aunque sea de forma breve, las diferencias existentes entre los términos *cuestionarios* e *inventarios*. Así, es posible encontrar autores que diferencian entre ambos tipos de procedimientos y otros que los consideran similares. Por ejemplo, FERNÁNDEZ BALLESTEROS (1992) sugiere que la diferencia entre ambos se establece en el tipo de respuesta que ofrece cada uno. De esta forma y según esta autora, los cuestionarios, generalmente, conllevan respuestas expresadas de forma nominal, mientras que los inventarios presentan posibilidades de respuesta nominal u ordinal. Habría que mencionar también las escalas de opinión o actitudes, las cuales tienen un formato de respuesta ordinal o de intervalo. Para otros autores, no obstante, la diferencia entre el cuestionario y el inventario reside, no tanto en la forma de responder, sino en la forma de preguntar. Así, para RODRÍGUEZ y BLANCO (1986), en los cuestionarios los ítems se formulan de forma interrogativa, y en los inventarios de forma asertiva.

Otros autores establecen otras diferencias entre ambos términos, por lo que no puede decirse que exista un consenso, más o menos generalizado, ante esta cuestión. En nuestra opinión, no existen diferencias claramente establecidas entre uno y otro procedimiento, por lo que nos mostramos de acuerdo con MATUD (1993) y consideramos que ambos términos se pueden utilizar de forma indistinta, aunque usaremos más a menudo el de «cuestionario», ya que es éste el que se utiliza más frecuentemente.

No obstante esta similitud, consideramos que cuando, tanto cuestionarios como inventarios, están diseñados según los criterios que permiten la medición de constructos

o características, pertenecen a otra categoría de instrumentos que en esta obra será objeto de estudio en el siguiente capítulo. En éste, nos dedicamos a un uso menos restringido de estos instrumentos, uso que conlleva que el proceso de diseño y construcción no sea tan riguroso como en el caso de los cuestionarios o inventarios estandarizados y comercializados habitualmente como los tests.

A la hora de clasificar los cuestionarios, existen diferentes criterios que se han utilizado para tal fin. MATUD (1993) resume algunas de las clasificaciones existentes en las siguientes:

— Según el *tipo de ítems*, cabría hablar de cuestionarios abiertos, cerrados, de elección múltiple, escalas de estimación, etc.

— Según la *escala de respuesta*, se pueden diferenciar entre los cuestionarios que buscan juicios categóricos y los que buscan juicios continuos (bien sea por estimación directa, o a través de los métodos comparativos).

— Según el *método de administración* (personal, por teléfono, correo, ordenador...).

— Según el *tipo de información* que suministran, se puede diferenciar entre cuestionarios de información general y cuestionarios específicos.

En general, los cuestionarios resultan una fuente de información muy útil en los procesos diagnósticos. Entre las ventajas que les podemos atribuir está el hecho de que son aplicables a multitud de variables y temas, además de a un conjunto amplio de sujetos. Asimismo, puesto que suministran información cuantificable, pueden puntuarse de forma objetiva, minimizando la inferencia por parte del evaluador. Por otra parte, su administración y análisis no requieren un excesivo coste —ni en términos económicos, ni en términos de esfuerzo y tiempo invertidos—. Pero, al igual que sucede con otros procedimientos, la información que suministran ha de ser complementada y cotejada con la proveniente de otras técnicas de evaluación, ya que existen diferentes fuentes de sesgos que pueden afectar a la calidad de la información recogida a través de ellos.

3.2. Diseño de cuestionarios y entrevistas estructuradas

Antes de adentrarnos en el proceso de preparación de los cuestionarios, recordamos al lector que, puesto que estos instrumentos comparten muchos rasgos en común con la entrevista estructurada, el contenido de este apartado puede aplicarse indistintamente a ambas técnicas, ya que, en momentos anteriores, hemos insistido en las diferencias básicas entre las mismas.

La realidad socio-educativa sobre la que revierte todo diagnóstico es compleja y su captación sumamente delicada, de tal forma que los medios que empleamos para la misma han de ser utilizados con el máximo cuidado y la mayor sabiduría. Evidentemente, esta afirmación puede aplicarse a la elaboración de los procedimientos

estructurados de encuesta (cuestionarios y entrevistas), ya que éstos se constituyen en instrumentos a usar en los procesos diagnósticos. Sobre este tema, dice MUCCHIELLI (1974:18): «quien prepare apresuradamente un cuestionario para hacerse una idea de un fenómeno (...) va de pesca sin demasiadas probabilidades de éxito».

Tanto el cuestionario como la entrevista cerrada son instrumentos estructurados y, como ya hemos insistido en numerosas ocasiones, requieren un detenido y riguroso proceso de diseño y construcción, que comienza con la elaboración de un marco conceptual, desde el cual se diseña el instrumento. Este principio no siempre es tenido en cuenta por los profesionales que se enfrentan por primera vez a la elaboración de un instrumento de este tipo. Así, en numerosas ocasiones hemos visto cómo la preparación de un cuestionario o entrevista comenzaba con la formulación de una lista de preguntas, tal cual éstas aparecían en la mente del constructor del mismo. Esta forma de actuar sería el equivalente a comenzar la construcción de un edificio por el tejado y las paredes. Es evidente que las preguntas son los «indicadores visibles» del cuestionario o la entrevista (como lo son las paredes y el tejado de una casa), pero estas preguntas no se sostienen sin una estructura que les da sentido y coherencia (los cimientos de todo edificio). Un cuestionario o entrevista elaborado sobre la base de un listado de preguntas se desmoronará, como un edificio sin cimientos, en el momento en que se someta a contraste con el «mundo real» (los encuestados que tienen que responder al instrumento). Por eso, estamos de acuerdo con MUCCHIELLI (1974:21-22) cuando afirma que «el tiempo invertido en la cuidadosa preparación de la encuesta antes de su ejecución efectiva (...) es un tiempo ganado en beneficio del valor de las informaciones y de las deducciones».

Es preciso hacer constar que siempre existe un marco teórico —explícito o no—, cuando se construye un instrumento. Incluso cuando, como en el ejemplo que hemos puesto, alguien se pone a formular preguntas sin orden ni concierto. Supongamos que, elaborando una encuesta sobre el ocio de los jóvenes, formulamos la siguiente pregunta: «¿Qué haces con tus amigos los fines de semana? a) Ir al cine; b) Ir a un pub o discoteca; c) Hacer botellones; d) Otros». El redactor de esta pregunta está partiendo de diversos supuestos implícitos; por ejemplo, considera que los jóvenes salen y no pasan su tiempo de ocio leyendo o viendo la televisión; está considerando también que salen en grupo de amigos y no solos; considera que salen los fines de semana; considera que las tres actividades más frecuentes entre la población de jóvenes son ir al cine, a la discoteca o hacer botellones y no otras como pueden ser viajar, pasear, salir de compras, practicar juegos de mesa o deportivos, etc.

Por lo tanto, el primer paso a dar en la elaboración de un instrumento cerrado de encuesta será hacer explícito el marco conceptual de supuestos previos desde el que se va a construir el instrumento. A partir de ahí se realizan diversas tareas, como veremos a continuación, que garantizarán la bondad psicométrica del instrumento finalmente diseñado. Sin embargo, incluso en un proceso lineal y operativo como es la construcción de encuestas estructuradas, la experiencia de la persona que la diseña es de gran ayuda

para conseguir un buen producto final. Y es que, como dice GARCÍA FERRANDO (1986:135), se aprende a construir estos instrumentos con la experiencia, sobre todo «con la cosecha de fracasos». Estos fracasos son los que nos ponen sobre aviso de las múltiples dificultades que cada pregunta o expresión concreta puede presentar cuando sea finalmente expuesta a la prueba de fuego: su aplicación a las personas concretas a las que el instrumento va dirigido.

3.2.1. Proceso de diseño y preparación de cuestionarios y entrevistas estructuradas

En este apartado pretendemos sintetizar las tareas básicas que requiere la construcción de un instrumento de encuesta estructurada. Se trata aquí de dar una visión general del proceso a seguir, proceso similar, por otra parte, al ya expuesto para la elaboración de sistemas categoriales de observación. Sin embargo, mientras que en la observación, la percepción de la realidad por parte del observador es el procedimiento básico para la recogida de información, la estrategia clave de los métodos basados en la encuesta es la *pregunta*. De forma genérica, podemos decir que elaborar y aplicar procedimientos de encuesta consiste en poner en práctica el *arte de preguntar*.

Por esta razón, además de la visión general del proceso a seguir en la elaboración de encuestas estructuradas, consideramos necesario dedicar un apartado a profundizar en la formulación y selección de las preguntas, ya que estas tareas concretas son de máxima importancia en la elaboración de un cuestionario o entrevista estructurada. Son tantos los tipos de preguntas posibles, los criterios a tener en cuenta para su formulación y las fuentes de sesgos en que pueden incurrir, que conviene dedicar apartados específicos a estas cuestiones que completarán y matizarán la visión general que en este apartado se ofrece.

Teniendo todo esto en cuenta, pasamos a describir el proceso de elaboración de encuestas estructuradas. En el anexo 4 aportamos un ejemplo de las tareas aquí apuntadas.

1. En primer lugar, será necesario *contextualizar el uso de la técnica en el proceso de diagnóstico*. Es evidente que un cuestionario o entrevista se usa para un fin que se inserta en el marco general de un proceso de diagnóstico en el que pueden incluirse además otros fines y otras técnicas de recogida de información. Por lo tanto, la construcción de estos instrumentos comenzará con la *formulación del problema* a explorar o los objetivos que con ellos se persiguen.

2. En segundo lugar, será necesario plantear y explicitar el *marco teórico-conceptual* desde el que parte el profesional del diagnóstico. Tanto el cuestionario como la entrevista estandarizada son técnicas estructuradas para la recogida de información, por lo que es absurdo e incoherente construirlas sin esa «estructura» que es el cimiento necesario para conseguir un buen instrumento.

   Esto implica definir los constructos o características con los que estamos trabajando. ¿Qué entendemos por ellos?, ¿qué contenidos engloban?, ¿qué

significado tienen en nuestro enfoque?... Para realizar satisfactoriamente esta tarea es imprescindible disponer de unos conocimientos previos. Estos conocimientos pueden provenir de anteriores experiencias y también de la lectura de la información teórica y empírica existente sobre el tema. Se trata, por tanto, de conocer qué enfoques respecto a nuestro problema existen, qué definiciones dan otros profesionales de los temas con los que vamos a trabajar, cuál es el contenido de estos temas, qué aspectos se han investigado con anterioridad, qué resultados ha arrojado la investigación sobre el tema, etc. Responder a éstas y otras preguntas es necesario para poder delimitar cuál es el modelo de referencia del que partimos (la «estructura» que dará sostén a nuestro instrumento). Puede resultar que nos adscribamos plenamente a las teorías y enfoques ya existentes o que tomemos parte de sus aportaciones para reconfigurar una nueva posición teórica al respecto: la nuestra propia, la cual será sometida a prueba a través del instrumento que vamos a usar. En la medida en que dicho instrumento demuestre ser válido, indirectamente nos informará de la «validez» que tiene también el modelo teórico diseñado.

Para no extendernos demasiado en estas cuestiones y ofrecer al lector una propuesta concreta de actuación en este segundo paso, a continuación sintetizamos las tareas que hay que realizar una vez concluida la revisión de aportaciones teórico-empíricas:

— Definir el *constructo*, característica o tema a diagnosticar. De qué conceptos partimos, cómo entendemos esa característica, qué elementos esenciales la componen... todo ello, lógicamente, referido a la población destinataria del diagnóstico. Así, no será lo mismo hablar de las «creencias sobre los anticonceptivos» en niños, adolescentes o personas adultas.

— Explicitar y definir las *dimensiones* de las que se compone el tema que estamos abordando. Por lo general, los temas que abordamos en un cuestionario y entrevista son complejos y, dentro de ellos, se engloban, a su vez, otros sub-temas relacionados entre sí y relacionados con el constructo general objeto de exploración. Aunque no siempre tiene por qué ser así. A menudo partimos de características que son unidimensionales, es decir, que no se pueden dividir en distintas partes diferenciadas porque todas tratan de lo mismo. El hecho de que un constructo sea uni o multidimensional dependerá evidentemente del fenómeno que estemos estudiando, pero también de la perspectiva que adopte el investigador. En el ejemplo del anexo 4 («Conducta y hábitos lectores en los adolescentes»), es fácil pensar que estamos ante un constructo multidimensional: hay muchas conductas, factores o aspectos en los que se manifiestan la conducta y los hábitos lectores. Así, habrá que prestar atención a varios aspectos que, aunque diferentes entre sí, guardan una relación: conocimiento y uso de la biblioteca, actitud ante la lectura, hábitos lectores, etc. El que un adolescente conozca y use la biblioteca no quiere decir que lea a menudo en casa o que considere que la lectura es una actividad útil y entretenida. Por eso, es más lógico

pensar que estamos ante varias dimensiones de una misma realidad. A pesar de ello, aunque bastante insostenible desde un punto teórico —y como demuestran resultados de muchas investigaciones—, un profesional del diagnóstico podría pensar que ir a la biblioteca o leer en casa o tener una pequeña biblioteca personal son conductas exactamente iguales entre sí, y que no existe ninguna diferencia teórica-empírica entre ellas. En ese caso, definiría una única dimensión que daría lugar, con posterioridad, a las preguntas del instrumento.

— Enumerar y definir los *indicadores* observables que operativizan el significado de cada dimensión definida. En la formulación de estos indicadores se deben seguir reglas similares a las utilizadas en la elaboración de sistemas de categorías; así, los indicadores deben ser definidos con claridad y precisión, ser mutuamente excluyentes, recoger los distintos aspectos o matices de la dimensión sin que falten elementos de la misma que no hayan sido operativizados en indicadores, etc. En la medida de lo posible, los indicadores se referirán a hechos, conductas y acciones concretas, que no conlleven ambigüedad y que puedan entenderse de la forma más objetiva posible.

3. Elaboración de un *banco amplio de posibles preguntas* a realizar. Estas preguntas serán redactadas a partir de los indicadores definidos en el paso anterior y deben recoger plenamente el contenido de cada uno de ellos. En este momento no conviene preocuparse demasiado por la calidad de las preguntas, sino por su cantidad. Es una fase creativa, de generación de multitud de preguntas, en la que conviene que participen diferentes personas a fin de obtener una mayor diversidad en los enunciados y alternativas de respuesta.

4. *Selección de las mejores preguntas* y reformulación de las mismas de acuerdo con los criterios al uso. Ahora sí que es apropiado seleccionar aquellas preguntas que consideramos que reúnen mejor los requisitos que veremos en apartados posteriores.

5. *Confección del primer borrador* o boceto del cuestionario (o de la cédula de la entrevista). Ello supone estudiar el orden en que se van a presentar las preguntas, además de preparar una presentación del instrumento en la que se indiquen las finalidades del diagnóstico, así como las instrucciones para la cumplimentación. Puesto que ya se han indicado los aspectos a tener en cuenta en la preparación de la cédula de la entrevista, en un apartado posterior haremos alusión a la preparación del boceto del cuestionario.

6. *Estudios piloto* de la versión inicial. Por mucho cuidado que se haya puesto en la confección del cuestionario o la entrevista, es necesario realizar algún tipo de estudio previo del instrumento antes de utilizarlo para la finalidad con la que fue diseñado. Una pregunta perfecta, a juicio del constructor del instrumento, puede presentar dificultades en su interpretación. De hecho, puede decirse que ningún instrumento es definitivo. Así, aunque haya sido aplicado en ocasiones anteriores,

cada vez que lo presentamos ante un colectivo de personas diferentes, se puede poner de manifiesto algún rasgo que denote su inadecuación. Es por ello que cuanto más completo y complejo sea el estudio piloto, más posibilidades habrá de que el instrumento esté más depurado. Se trata, por tanto, de poner diferentes «filtros» que vayan contribuyendo a la mejora continua de nuestro cuestionario o entrevista, aunque esto dependerá también del alcance del proceso de diagnóstico y de las condiciones de partida con las que trabajamos. Existen diversas posibilidades a la hora de realizar estudios previos de un instrumento, aunque lo deseable es combinar varias de estas opciones con la finalidad de conseguir un resultado mejor:

— *Revisión por parte de especialistas* tanto en la construcción de instrumentos, como en la temática abordada por el cuestionario o entrevista. La experiencia de estas personas puede ayudarnos a encontrar defectos técnicos y puntos oscuros en el instrumento que hemos diseñado. A los especialistas podemos pedirles que señalen todas las posibles dificultades que encuentren, que sugieran formas alternativas de redacción o presentación y que valoren numéricamente la calidad de las preguntas u otros elementos del instrumento.

— *Revisión por parte de los sujetos*. En este caso, son sujetos con características similares a los que compondrán la muestra de estudio los que opinan sobre el instrumento, indicando las dificultades que han tenido para interpretar y comprender las preguntas. No se trata tanto de que cumplimenten la prueba, como de que comenten en voz alta las observaciones que consideren oportunas. Aunque se trate de un cuestionario, es preferible realizar esta prueba bajo el formato de una entrevista, con la presencia del profesional que elabora el instrumento, quien podrá observar a los sujetos y comprobar de primera mano las dificultades existentes en el mismo.

— *Prueba piloto*, propiamente dicha. En este caso, sujetos de las mismas características de las personas a las que va dirigido el cuestionario o entrevista cumplimentan el instrumento, y sus respuestas son sometidas a análisis estadísticos que nos permiten detectar los defectos en su construcción. Así, un simple estudio de porcentajes, medias y desviaciones típicas puede decirnos mucho sobre las preguntas formuladas y sus alternativas de respuesta. Algunos de los aspectos en que debemos fijarnos son: tendencias de respuesta diferentes a las que se esperan, excesivo número de sujetos que no responden a un ítem dado (ambas circunstancias pueden indicar que se ha formulado mal una pregunta, o que ésta no ha sido entendida), alternativas de respuestas que concentran un elevado porcentaje, etc.

La forma más adecuada de hacer este estudio piloto es seleccionar a un grupo reducido de personas (entre 50 y 60) en el que estén representados los diversos sub-grupos a los que nos dirigimos (hombres y mujeres, personas de todos los grupos de edad, etc.). Si este criterio no se tiene en cuenta, la prueba piloto pierde su utilidad. Por

ejemplo, supongamos que en un proceso de evaluación en el que se está trabajando con personas de todas las edades, no seleccionamos en la muestra piloto a sujetos con edad superior a 50 años. Es posible que si, terminada la prueba piloto, aplicamos el cuestionario a sujetos de estas edades, éstos detecten en el instrumento errores o defectos que podían haber sido corregidos con anterioridad y que, sin embargo, no han aparecido como significativos ante personas de otras edades.

Es importante que el estudio piloto (así como las primeras administraciones del estudio definitivo) sea realizado por el constructor del instrumento, ya que constituye una oportunidad excelente para detectar posibles deficiencias en el mismo. Por las reacciones y preguntas de los encuestados podemos detectar qué es lo que no funciona o inferir las dificultades que las personas encuentran a la hora de responder. Asimismo, KORNHAUSER y SHEATSLEY (1976) sugieren que si se llevan a cabo cambios sustanciales en el instrumento como consecuencia de la prueba piloto, es necesario llevar a cabo un segundo pretest (e incluso un tercero o cuarto, si así lo requiere la situación). Esto es lo que se denomina como *procedimiento en cascada*.

3.2.2. Las preguntas en cuestionarios y entrevistas estructuradas: tipos, selección y formulación

Dado que la formulación de preguntas en las técnicas de encuesta constituye la estrategia básica a partir de la cual se recopila la información, vamos a dedicar un apartado a analizar con más profundidad esta cuestión. Así, en primer lugar, haremos referencia a los distintos tipos de preguntas a utilizar en estos instrumentos, para pasar después a revisar los diversos aspectos que hay que tener en cuenta en su formulación y selección.

Los posibles tipos de preguntas a utilizar en los cuestionarios no tienen, en principio, por qué ser diferentes a los usados en las entrevistas estructuradas, por lo que todo lo que aquí apuntamos sobre esta cuestión es aplicable a ambas técnicas de encuesta.

En primer lugar, en la pregunta cabe diferenciar dos partes: el enunciado que constituye la pregunta en sí misma, y las alternativas, que son las posibles respuestas entre las cuales el sujeto elegirá una (o varias, si es el caso). En las preguntas abiertas, que veremos a continuación, no se ofrecen alternativas de respuesta, sino que la pregunta está constituida exclusivamente por el enunciado. De cualquier forma, para ciertos fines, como ocurre en algunos estudios de mercado, los enunciados y alternativas adoptan formas diferentes a las aquí expuestas (por ejemplo si se da a probar un producto para ver si gusta o no, si se presentan fotografías para que seleccione una imagen, etc.).

Autores como FOX (1981) consideran que en cierto sentido sólo existen dos tipos de respuestas a las preguntas de una encuesta: la respuesta libre, que otorga libertad al sujeto para determinar su respuesta, y la respuesta estructurada, en la que se sugieren algunas o todas las respuestas potenciales para que el sujeto elija una o varias de ellas. Nosotros seguiremos esta clasificación, incluyendo en ella diferentes tipos de respuestas estructuradas o cerradas.

En las primeras, *preguntas abiertas o de respuesta libre* (ver cuadro 1), no se prevén las respuestas a dar y se concede libertad al sujeto para expresarse en los términos que considere oportuno.

CUADRO 1

EJEMPLOS DE PREGUNTA ABIERTA

- ¿Qué actividades practica en su tiempo libre?
- ¿Cuáles son las razones por las que asiste a este centro?
- ¿Qué asignaturas cursó en la Universidad?
- ¿Cuál es el último libro que ha leído?

No obstante, algunas preguntas en apariencia abiertas, son en realidad cerradas. Por ejemplo: «¿Ve Vd. la televisión?». Aunque, materialmente, no le damos dos alternativas a escoger, a la hora de responder la persona entrevistada no tiene más opción que contestar con un «sí» o con un «no». Cualquier respuesta diferente a ésta que dé, o bien será categorizada por el entrevistador en alguna de las dos modalidades enunciadas (por ejemplo, si contestara «la veo un par de horas todos los días», el entrevistador registrará un «sí»), o bien será desestimada como respuesta o asignada a la categoría «no sabe/no contesta» (por ejemplo, si responde «¿y a Vd. qué le importa?», o «mi marido sí que ve mucho la tele»).

Entre las ventajas de este tipo de preguntas, cabe resaltar que proporcionan información única, al tiempo que permiten abordar a cualquier sujeto y obtener de él informaciones verdaderamente útiles que pueden ser diferentes a las que considera inicialmente el encuestador. Sin embargo, plantean el problema de que resultan más complejas y dificultosas a la hora del análisis, ya que cada persona contestará de forma diferente y alegará razones o contenidos diferentes, por lo que este tipo de preguntas requiere un análisis de contenido que permita establecer cuáles son las categorías de contenido que aglutinan a toda la diversidad de respuestas ofrecidas por los encuestados.

Una opción diferente a la pregunta abierta consiste en indicarle al sujeto qué tiene que responder, de manera que no «elabora» su respuesta, sino que «elige», entre las que se le ofrecen, la que considera más adecuada según la petición formulada en el enunciado de la cuestión. Este tipo de preguntas es el que se denomina como *cerrada* o *estructurada*; su finalidad, generalmente, es solicitar respuestas del tipo «aprobación-desaprobación» o de selección entre un conjunto de opciones previamente establecidas. Dentro del grupo

de preguntas cerradas están las que son *cerradas dicotómicas* (ver cuadro 2), en las que sólo se plantean dos opciones de respuesta, generalmente una de ellas afirmativa y la otra desestimativa o negativa.

CUADRO 2

EJEMPLOS DE PREGUNTA CERRADA DICOTÓMICA

- ¿Ha trabajado antes para esta empresa?
- ¿Lee usted el periódico?
- ¿Tiene hijos?
- ¿Cursó el último año de carrera?

Este tipo de preguntas debe utilizarse con ciertas precauciones en el campo educativo. En ocasiones, las actitudes y las conductas no pueden categorizarse de forma dicotómica, ya que toman la forma de un continuo. Si tomamos como ejemplo la pregunta «¿lee Vd. el periódico?», nos puede contestar «sí» un sujeto que lo lee a diario, un sujeto que lo lee el domingo, y también un sujeto que lo lee dos veces al año. Es evidente que estas situaciones son bien distintas, por lo que este tipo de pregunta no permite discriminar la respuesta, forzando una generalización extrema. Esto sería justificable tan sólo si tenemos la intención de diferenciar a los sujetos en dos grupos, los que leen el periódico y los que no lo leen, pero aun en ese caso, corremos el riesgo de que prácticamente todo el mundo afirme leerlo. Es por ello que este tipo de cuestiones sólo debe utilizarse en los casos en que la elección sea realmente bidimensional; es decir, en los casos en los que los sujetos tengan que adscribirse a una u otra modalidad, sin que existan opciones intermedias.

Otra modalidad de preguntas que también proporciona las respuestas de antemano al encuestado, aunque en este caso permitiéndole elegir entre un conjunto mayor de opciones, son las preguntas de *elección múltiple*. En este caso, podemos encontrar dos formas diferentes de plantear estas preguntas: por un lado, cuando el entrevistado tiene que elegir entre opciones de igual «valor», es decir, entre las diversas opciones presentadas, no existe una relación de intensidad o cantidad. Este tipo de preguntas se denominan de *elección múltiple en abanico* (ver cuadro 3).

Para la correcta formulación de estas preguntas es necesario conocer bien las tendencias en la población a la que se va a aplicar la encuesta, a fin de que las opciones que se les ofrecen a los sujetos sean las más frecuentes entre dicha población. De cualquier forma, es conveniente incluir en ellas una modalidad de respuesta denominada

«otra», de tal manera que se puedan adscribir a ella los sujetos que no consideran adecuada ninguna de las opciones que se ofrecen. Como vemos en el cuadro 3, en la modalidad «otra» puede dejarse o no un espacio para que el sujeto consigne la respuesta que no está incluida en las opciones ofrecidas.

CUADRO 3

EJEMPLOS DE PREGUNTA DE ELECCIÓN MÚLTIPLE EN ABANICO

➤ **¿Cuál es la principal dificultad que encuentra a la hora de asistir a las clases de educación de personas adultas?**

— La falta de tiempo libre
— La oposición de mi familia
— La dificultad de entender lo que se explica en clase
— La falta de motivación para asistir a las clases
— Otra. Indicar: _____

➤ **Señale, entre las siguientes, la razón por la que solicita este empleo:**

— El prestigio de la empresa
— El tipo de trabajo a realizar
— Las posibilidades de ascenso que ofrece
— La remuneración del puesto
— Otra

Cuando el cuestionario o entrevista está en la fase de estudio piloto, es conveniente que en la modalidad «otras», los sujetos puedan indicar su respuesta. De esta forma, si tras los análisis estadísticos se observa que un porcentaje escaso de sujetos ha seleccionado las modalidades cerradas, mientras que un elevado porcentaje ha seleccionado la opción «otra», es evidente que las alternativas de respuesta formuladas no han sido las más adecuadas, por lo que la pregunta habrá de ser mejorada para futuras aplicaciones. Para conocer cuáles son las respuestas más frecuentes entre la población serán de gran utilidad las razones o informaciones aportadas por los sujetos en la modalidad «otra».

Por otra parte, están las preguntas entre cuyas opciones de respuesta sí existe una gradación o relación de intensidad o cuantía. Éstas son las preguntas denominadas de *elección múltiple de estimación* (ver cuadro 4). Ante ellas el sujeto debe indicar el grado que le parece que mejor representa lo que el enunciado le ha solicitado valorar. Es habitual utilizar la escala tipo *Likert*, en la cual se expresa la opinión estimando el nivel de acuerdo en una escala de puntos que oscila entre el total desacuerdo y el total acuerdo ante la afirmación o enunciado que se presenta.

CUADRO 4

EJEMPLOS DE PREGUNTA DE ELECCIÓN MÚLTIPLE DE ESTIMACIÓN

➤ **¿Cuál es su opinión general respecto al convenio colectivo de su empresa?**

— Totalmente de acuerdo
— Bastante de acuerdo
— Indiferente
— Bastante en desacuerdo
— Totalmente en desacuerdo

➤ **¿Con qué frecuencia visita la biblioteca pública?**

— Menos de una vez al mes
— De una a dos veces al mes
— De tres a cinco veces al mes
— Más de cinco veces al mes

Siempre que resulte posible, es preferible usar alternativas de respuesta como las del segundo ejemplo (menos de una vez al mes, de una a dos veces...). Expresiones como «siempre», «a veces»... o «totalmente de acuerdo», «bastante de acuerdo»... resultan demasiado ambiguas, ya que, ante una misma «cantidad» de la conducta, un sujeto puede tender a calificarla de «mucha» y otro sujeto de «bastante» o incluso «alguna».

Por otra parte, cuando se usan este tipo de escalas hay que tener en cuenta un conjunto de cuestiones que afectan a la precisión de las respuestas y a los posibles sesgos en las mismas. Un aspecto que tiene en ello una gran incidencia es el número de alternativas de respuesta que se presenta, lo que implica la capacidad de discriminación de la escala y su fiabilidad. Así, a medida que disminuye el número de categorías utilizado, la fiabilidad parece ser menor, aunque resulta inapreciable la pérdida de fiabilidad entre utilizar 7 ó 10 categorías y por ello se suele recomendar un mínimo de estimaciones entre 5 y 7 puntos (MATUD, 1993).

Entre las ventajas de los distintos tipos de preguntas cerradas, podemos mencionar las siguientes (MATUD, 1993; MUCCHIELLI, 1974; SELLTIZ et al., 1976):

— Permiten clasificar rápidamente cualquier respuesta, resultan sencillas y rápidas de administrar y fáciles de cuantificar para realizar los correspondientes análisis estadísticos.

— Nos permiten asegurarnos de que las respuestas se dan desde el marco de referencia que es significativo para el profesional del diagnóstico. Si preguntamos «¿a cuántos cursos de formación asiste Vd.?», nos vamos a encontrar quien nos responde a la pregunta teniendo en cuenta toda su vida («he asistido a cinco cursos

en total»), quien lo hace refiriéndose al último año («durante este año sólo he ido a uno»), quien lo hace por meses o, incluso, quien nos hace una lista detallada de todos los cursos a los que ha asistido o quien nos dice simplemente que «muchos» o «pocos». Por eso, las preguntas cerradas ofrecen un marco homogéneo para la respuesta de los sujetos. Todos contestan en los mismos términos.

— Relacionado con lo anterior, en ocasiones la provisión de alternativas de respuestas contribuye a aclarar y matizar el significado de la pregunta.

— Exigen un menor esfuerzo al entrevistado, en tanto que éste se limita a señalar la opción de respuesta y no tiene que pensarla, elaborarla y comunicarla (lo que puede ser de gran utilidad para fomentar su respuesta y la consecuente devolución del cuestionario, si se trata de esta técnica).

El principal inconveniente que de ellas deriva es que el sujeto no puede expresar su propia opinión con sus propias palabras, por lo que las opciones de respuesta que le ofrece el entrevistador pueden parecerles poco adecuadas o incompletas para representar su opinión. No obstante, como se verá en un momento posterior, esto sucederá en mayor medida si la pregunta y sus opciones de respuestas no están bien formuladas. Como ya hemos planteado, dejar una alternativa de final abierto es una opción recomendable para aquel sujeto que considera que su opinión no está reflejada en el abanico de respuestas ofrecido.

Con o sin alternativa abierta, es esencial que las alternativas de respuesta que figuran en el instrumento hayan sido cuidadosamente pensadas, diseñadas e incluidas, como fruto de un proceso de estudio y análisis previo del investigador cuando define y realiza el marco conceptual en el que basa su instrumento. Una mayoría de los sujetos se limitará a seleccionar una alternativa entre las que se le ofrecen, por lo que la omisión de alternativas relevantes puede alterar considerablemente la respuesta.

Asimismo, otra dificultad inherente al uso de preguntas cerradas estriba en que puede llegar a forzar un juicio de opinión en un tema en el que el sujeto todavía no ha cristalizado una opinión. Para ello, en determinadas ocasiones, es adecuado incluir la alternativa «no sé», a fin de que el sujeto perciba que no tener formada una opinión al respecto también es una alternativa posible y aceptable para el encuestador (SELLTIZ et al., 1976).

Por otro lado, es necesario resaltar los inconvenientes concretos que presentan las preguntas de estimación. Como se indicó en el capítulo dedicado a la observación, ante este tipo de escalas, es de destacar la especial incidencia de ciertos sesgos que pueden afectar a la respuesta. Entre ellos, el *efecto de halo* (resulta posible calificar los ítems rápidamente de forma global y sin prestar la necesaria atención a cada enunciado concreto), y los errores de tendencia central, de lenidad y severidad.

Todos los tipos de preguntas presentados difieren entre sí por la *forma* que adoptan. Sin embargo, las preguntas también se pueden clasificar de acuerdo con otros criterios

como su *contenido*, como hace PATTON (1990), quien distingue entre preguntas sobre experiencias y comportamientos, sobre opiniones y valores, sobre sentimientos y emociones, sobre conocimientos y sobre cuestiones sensoriales. PATTON (1990) señala también que las preguntas pueden ser de tipo biográfico-demográfico. Este tipo de preguntas se enfoca a conocer determinadas características de los sujetos (su edad, nivel de estudio, sexo...) y, generalmente, aportan información que se va a utilizar para establecer contrastes y diferencias. Es decir, si preguntamos a una persona su edad, probablemente esa información no sea directamente útil desde el punto de vista de los objetivos de nuestro diagnóstico, sino que va a ser usada para conocer las diferencias en las opiniones, conductas, actitudes, etc. de los sujetos en función del grupo de edad a que pertenecen.

Las preguntas pueden también clasificarse según la *función* que cumplen dentro de la entrevista. Utilizando este criterio, CABRERA y ESPIN (1986) distinguen entre los siguientes tipos de preguntas: de consistencia (para asegurarse de la sinceridad del encuestado), de introducción o «rompehielos», amortiguadoras (formuladas de tal forma que reducen la brusquedad o rudeza de su contenido), de batería (preguntas sobre una misma cuestión que se complementan), y de filtro (se realizan previamente a otra pregunta a fin de eliminar a los sujetos a los que no afecte su contenido).

Una vez descritos los diversos tipos de preguntas a usar en las técnicas de encuesta, pasamos ya a exponer los criterios que hay que tener en cuenta en su formulación. Estos criterios permitirán que el instrumento finalmente diseñado no presente defectos importantes que condicionarán en gran medida la calidad de la información recogida. En general, siempre que se formulan preguntas para un instrumento como el cuestionario o la entrevista estructurada, hay que cuidar:

— El lenguaje que se usa. Éste debe ser totalmente comprensible para la persona o personas encuestadas. A veces esta tarea puede resultar compleja. Si la población destinataria de la entrevista es muy diversa en cuanto a su nivel socio-cultural, lo que para unos puede resultar un lenguaje «comprensible», para otros puede suponer una simplificación excesiva. En todo caso, debe evitarse el uso de un lenguaje vulgar y malsonante que perjudique la imagen del instrumento y del profesional que lo usa.

— La claridad de su formulación, de tal forma que la intención de la pregunta y la naturaleza de la información que se busca deben estar claras para el encuestado.

— La concreción del contenido y del período de tiempo. Aunque se relaciona mucho con lo anterior, una pregunta puede estar expresada en lenguaje claro y sin embargo, aparecer como ambigua y confusa la intención del encuestador. El ejemplo que pone FOX (1981) es el siguiente: «¿Cuál ha sido su experiencia más interesante desde que se hizo profesor?». Tal y como plantea este autor, el período de tiempo que abarca la pregunta es razonablemente claro, sin embargo, los límites de su contenido no lo son. ¿Tiene que referirse el encuestado sólo a sus

experiencias docentes o puede hablar de «otras experiencias interesantes» desde que se hizo profesor (su boda, el nacimiento de su primer hijo...)?

— La unicidad del propósito, o asegurar que cada pregunta busca un elemento o fragmento de información y sólo uno. Lo contrario a esto es lo que se llama pregunta de doble-cañón, que son aquellas en las que se pregunta por más de una persona o cosa a la vez. Por ejemplo, «¿está Vd. de acuerdo con el convenio colectivo de la empresa y la forma de llevarlo a cabo por sus superiores?». Es evidente que si el entrevistado dice «sí» o «no», no podemos estar seguros de qué quiere decir. ¿Se refiere a que está de acuerdo o en desacuerdo con las dos cosas (el convenio y la forma de llevarlo a cabo), con una sola, con ninguna?

— Eludir los contenidos, palabras o expresiones que pueden tener una carga emocional muy fuerte. Ante estas preguntas, lo normal es encontrar respuestas estereotipadas que no aportan ninguna información. Por ejemplo, «¿Es partidario del terrorismo?».

— La existencia de una opción de respuesta cualquiera que sea la opinión del encuestado. Es decir, a no ser que se haya dispuesto previamente una pregunta filtro, es necesario que todos los sujetos puedan contestar con una aproximación razonable a su opinión o experiencia. Así, las alternativas de respuesta deben recoger todas las posibles opciones existentes a la hora de contestar a la pregunta. Esto implica que en las preguntas cerradas de elección en abanico, es necesario incluir la opción «otro» para que un sujeto pueda elegirla en caso de que no le satisfagan ninguna de las alternativas dadas.

— La independencia de sugerencias. En la formulación de la pregunta no debe haber nada que sugiera al sujeto que se espera de él cierta respuesta, o que ciertas respuestas son más deseables que otras.

— Asegurarse de que las preguntas dan las indicaciones completas para que el encuestado sepa cómo debe ser la naturaleza de su respuesta, y elaborar preguntas y alternativas de respuestas gramaticalmente coherentes.

COHEN y MANION (1990) aportan otras recomendaciones de gran interés. Utilizaremos también los ejemplos aportados por estos autores que, aparte de ilustrativos resultan divertidos. Así, los autores citados recomiendan evitar las *preguntas excesivamente complejas*, del tipo:

> «¿Preferiría un curso corto sin titulación (3, 4 ó 5 sesiones) con parte del tiempo libre (por ejemplo, las tardes de los miércoles) y una tarde a la semana de asistencia con gratificación en metálico por el viaje, o un curso más largo sin titulación (6, 7 u 8 semanas) con todo el tiempo libre, o todo el curso diseñado sobre parte del tiempo libre sin asistencia nocturna?» (COHEN y MANION, 1990:145).

Y las *preguntas o instrucciones irritantes*, como:

> «¿Ha asistido alguna vez en tiempo de servicio a un curso de cualquier clase durante su carrera entera de maestro? Si tiene más de 40 años y nunca ha asistido a un curso en la carrera, ponga una marca en la casilla rotulada NUNCA y otra en la casilla rotulada VIEJO». (COHEN y MANION, 1990:145).

Asimismo, hay que evitar las preguntas que emplean *dobles negaciones*, tales como:

> «¿Cuál es su sincera opinión sobre que ningún maestro pudiera apuntarse al curso con titulación y de carrera si no ha completado al menos dos años enteros de enseñanza? (COHEN y MANION, 1990:145).

Pese a lo obvio que pueda parecer que estas preguntas son inadecuadas, nuestra experiencia nos dice que son ejemplos que pueden darse en la realidad. En ocasiones, estamos tan centrados en «meter» en la pregunta toda la información que nos parece relevante, que el producto final es ininteligible para cualquier otra persona que no sea la que elabora la pregunta.

Además de todas las recomendaciones anteriores, otros aspectos han de cuidarse en la formulación y selección de las preguntas (GARCÍA JIMÉNEZ, 1997). Así, no se deben incluir preguntas cuya información pueda obtenerse por otra vía (a no ser que exista una intención definida de someter algo a contraste). Por ejemplo, si entrevistamos a profesores, debemos abstenernos de preguntarles por cuestiones como los objetivos didácticos, los contenidos, la metodología, etc., pues esta información puede obtenerse de sus programaciones curriculares, sin necesidad de hacer perder el tiempo a los profesionales que estamos entrevistando. Asimismo, en este tipo de instrumentos no conviene incluir preguntas confidenciales, ni tampoco preguntas cuya respuesta exija un gran esfuerzo o revistan especial dificultad para los entrevistados. Hay que recordar que al tratarse de técnicas estructuradas, la información que persiguen es inmediata y no requiere una elaboración de la respuesta como sucede en las entrevistas abiertas.

Completamos esta indicación del proceso a seguir en la formulación y selección de preguntas en los cuestionarios y entrevistas con un breve resumen de las preguntas recogidas en la excelente *Guía para la construcción de cuestionarios* de KORNHAUSER y SHEATSLEY (1976:608-632), cuya lectura detenida recomendamos a aquellos profesionales que desean diseñar este tipo de instrumentos. En el cuadro 5 se recogen las ideas de esta guía que consideramos más ajustadas al tipo de cuestionario/entrevista que aquí hemos presentado. Hay que advertir también que en el citado cuadro sólo se enumeran los aspectos que estos autores recomiendan comprobar en la formulación de las preguntas (no los comentarios sobre los mismos ni los ilustrativos ejemplos aportados por los autores).

3.2.3. Tendencias de respuesta que inducen al sesgo

Además de los conocidos efectos de halo, error de tendencia central, etc. en los cuestionarios y entrevistas estructuradas suelen incidir un conjunto de mecanismos de defensa del sujeto que conllevan la distorsión de la información que se recoge a través de estos medios. Algunos de estos fenómenos son recogidos por MUCCHIELLI (1974), quien los clasifica en dos grupos: por un lado, deformaciones que proceden de las defensas sociales automáticas de los encuestados y, por otro, deformaciones procedentes de la presentación y organización interna del instrumento. A continuación, pasamos a describir cada uno de estos tipos de deformaciones.

CUADRO 5

GUÍA PARA LA CONSTRUCCIÓN DE CUESTIONARIOS

➤ **Decisiones sobre el contenido de la pregunta**

- ¿Es precisa esta pregunta?, ¿cómo sería útil y hasta qué punto? (lo que implica cuestionar: ¿se halla cubierto ya suficientemente el tema por otras preguntas?, ¿requiere el asunto por sí mismo una pregunta separada, o puede ser integrada con otras preguntas?...)
- ¿Son precisas varias preguntas similares? (lo que implica cuestionar: ¿cubre la pregunta adecuadamente el área pretendida?, ¿debería subdividirse la pregunta?...)
- ¿Poseen los encuestados la información necesaria para contestar a las preguntas? (lo que implica cuestionar: ¿se trata de un asunto del que pueden informar adecuadamente?, ¿se halla el asunto dentro de la experiencia del encuestado?, ¿proporcionarán juicios adecuados y exentos de prejuicios?...)
- ¿Precisa la pregunta ser más concreta, específica y estrechamente relacionada con la experiencia personal del interrogado? (lo que implica cuestionar: ¿está formulada en una forma demasiado general?, ¿puede la información ser obtenida de un modo más sencillo mediante la referencia a la conducta del encuestado?...)
- ¿Se halla el contenido de la pregunta desviado o dirigido en una dirección? (lo que implica cuestionar: ¿sería igualmente aceptable el contenido para personas con opiniones opuestas?, ¿existe la posibilidad de obtener respuestas que favorecieran desmesuradamente un lado de la cuestión?...)
- ¿Darán los encuestados la información que se les solicita? (lo que implica cuestionar: ¿es la información que se solicita demasiado privada o de naturaleza embarazosa o de otro tipo que suele llevar a la resistencia o evasión?, ¿qué objeción puede tener la persona para contestar?, ¿muestra la pregunta tendencia a hallar influencias emocionales y deseos que llevarán a una falsificación de las respuestas?...)

➤ **Decisiones con respecto a la redacción de la pregunta**

- ¿Puede la pregunta ser entendida erróneamente?, ¿contiene una fraseología difícil o nada clara? (lo que implica cuestionar: ¿son las palabras lo suficientemente sencillas para ser entendidas por todos los encuestados?, ¿se usan términos especializados?, ¿es la estructura de la frase corta y sencilla?, ¿existe alguna ambigüedad o confusión?...)
- ¿Expresa la pregunta de forma adecuada todas las alternativas en relación con el tema?
- ¿Es claro y uniforme el marco de referencia para todos los encuestados?
- ¿Está afectada la redacción de algún error sistemático?, ¿se halla emocionalmente cargada o inclinada hacia una determinada clase de respuesta? (lo que implica cuestionar: ¿se utilizan tópicos?, ¿contiene nombres que llevan implícitos un cierto prestigio?, ¿utiliza términos superlativos que pueden inclinar la respuesta en un sentido?...)
- ¿Puede provocar algún rechazo o actitud defensiva la redacción de la pregunta?
- ¿Se obtendrían mejores resultados con una redacción más personal o bien menos personal de las preguntas?
- ¿Puede formularse mejor la pregunta en forma directa o indirecta?

➤ **Decisiones sobre las formas de respuesta a la pregunta**

- ¿Puede plantearse la pregunta con diversas formas (abierta, cerrada, cerrada con opción abierta)?

- Si se utiliza el tipo de reconocimiento, ¿cuál es la mejor forma (dicotomía, elección múltiple en abanico o de estimación)?
- Si se utiliza una lista (preguntas de elección múltiple en abanico), ¿cubre ésta todas las alternativas significativas de forma adecuada sin omisiones y en un orden lógico?, ¿tiene una longitud adecuada?, ¿es la redacción de las alternativas imparcial y compensada?
- ¿Reúne la forma de pregunta seleccionada las cualidades de fácil, definida, uniforme y adecuada para el propósito que se persigue?

➤ **Decisiones acerca de la posición de la pregunta en secuencia**

- ¿Tiende a ser influida la respuesta por el contenido de las preguntas precedentes? (lo cual implica cuestionar: ¿crean las preguntas anteriores un cierto ambiente o expectación que podría influir en las respuestas a esta pregunta concreta?, ¿resulta inadecuada esta pregunta si se hubieran emitido determinadas contestaciones anteriormente?...)
- ¿Sigue la secuencia de preguntas un orden psicológico correcto?, ¿se formulan en una secuencia natural?
- ¿Aparece la pregunta demasiado pronto o demasiado tarde desde el punto de vista del interés, la atención, etc.?

(Basado en KORNHAUSER y SHEATSLEY, 1976: 608-632)

Las deformaciones que proceden de las defensas automáticas de los sujetos pueden ser:

— La reacción de prestigio o «defensa de fachada», que se trata de una reacción automática correspondiente al miedo a ser mal juzgado a través de la respuesta. Es una tendencia de «fachada» en el sentido de que se trata de «guardar las apariencias». En ocasiones se manifiesta en la tendencia a situarse en un estereotipo o acatar los moldes socialmente establecidos. Por ejemplo, si preguntamos a alguien por qué lee, es más fácil que nos responda que lo hace «para adquirir nuevos conocimientos», a que diga que lo hace «para matar el aburrimiento».

— La contracción defensiva ante la pregunta personalizada. Así, en preguntas que solicitan de forma directa una opinión personal («¿qué piensa de...?, ¿cuál es su opinión sobre...?»), se corre el riesgo de suscitar respuestas evasivas, sobre todo, cuando no está clara la utilización que se va a hacer de la información.

— El temor a ciertas palabras con una alta carga afectiva o un valor socialmente no deseable. Ejemplos de estas expresiones pueden ser «castigo», «guerra», etc.

— Las referencias a personas concretas en las preguntas conlleva el riesgo de que las simpatías o rechazos que las mismas suscitan se transmitan al contenido de la

pregunta.

— El temor al cambio y la tendencia al conformismo. Si la pregunta se plantea de forma que haya una referencia explícita a un cambio susceptible de repercutir sobre el equilibrio existente, las reacciones ante el mismo pueden ser contrarias. Por ejemplo, ante la pregunta «¿considera que el gobierno debe cambiar la ley para realizar intercambios comerciales con los países árabes?», es más fácil encontrar rechazo que si la planteamos del siguiente modo: «¿considera que deberían realizarse intercambios comerciales con los países árabes?». En el primer ejemplo, la pregunta puede suscitar reacciones negativas que no tienen que ver con el contenido sobre el que se busca información.

Por otro lado, en la clasificación de MUCCHIELLI (1974) se recoge un segundo conjunto de factores que pueden distorsionar la información procedente de los cuestionarios y entrevistas estructuradas. Son las ya mencionadas deformaciones involuntarias procedentes de la presentación y organización interna del instrumento y se refieren a:

— La reticencia a concentrarse en el cuestionario por cuestiones como la exigencia de atención, la disponibilidad de tiempo, etc.

— La reticencia a los cambios bruscos de temas dentro del instrumento.

— El efecto de halo o contaminación o interferencia de unas preguntas con otras.

— Los efectos derivados de la longitud del cuestionario y de las preguntas. Las preguntas largas, complicadas, que exigen reflexión y atención provocan la tendencia a «declinar». Si, por otra parte, el cuestionario es demasiado largo, también se corre el riesgo de que el sujeto no lo cumplimente.

3.2.4. Preparación del formato del cuestionario

Como se ha visto anteriormente, la elaboración del formato del cuestionario —al igual que sucede con la cédula de la entrevista estructurada— es una tarea que adquiere sentido cuando se han realizado previamente ciertas operaciones relacionadas con la definición del constructo y sus indicadores, y la formulación y selección de las preguntas. Como en la entrevista, el cuestionario debe tener una fórmula de presentación e introducción. En el caso del cuestionario, dicha presentación es incluso más importante. No olvidemos que el encuestador no está aquí presente, por lo que es fácil que el sujeto no se sienta «obligado» o «comprometido» a responder. Por lo tanto, en este instrumento es fundamental que, en lugar bien visible, aparezca información sobre quién realiza la encuesta y con qué fin. Esto es lo que se denomina la introducción del cuestionario.

La introducción debe ser clara y breve a un mismo tiempo, ya que su longitud disminuye el efecto del impacto y aburre al lector (MUCCHIELLI, 1974). Básicamente, debe hacer referencia a las siguientes informaciones:

— En primer lugar, el cuestionario debe tener un título en el que se resuma de forma breve su contenido («cuestionario para la detección de los conocimientos previos de estadística», «cuestionario sobre la igualdad de oportunidades en el entorno laboral»...). Es conveniente también que, bajo el título, figure el nombre de su autor o autores.

— Asimismo, es de gran importancia indicar quién realiza el diagnóstico o investigación para el que se solicita la participación en el cuestionario. Un cuestionario que no indica por quién está realizado puede generar recelos y desconfianzas en el encuestado al desconocer éste quién va a hacer uso de la información que aporta.

— Muy en relación con lo anterior, es necesario indicar el fin o fines que se persiguen con el cuestionario o con el estudio en general. Gracias a este dato, no sólo es posible motivar al encuestado, que así entiende el por qué es necesaria su contribución, sino que, igual que en el punto anterior, se le aclara para qué sirve la información que aporta y se despejan posibles actitudes recelosas. Si el ámbito del diagnóstico es más amplio y complejo que el del cuestionario, se deben indicar ambos. Es decir, la finalidad general del diagnóstico (por ejemplo, conocer las causas que provocan el absentismo en una empresa) y el objetivo para el cual se usa el cuestionario (por ejemplo, determinar el nivel de satisfacción de los trabajadores respecto del ambiente en que trabajan).

— La garantía del anonimato de las respuestas es también una cuestión que debe ser explicitada en la introducción. El encuestado debe saber que en ningún momento su opinión será identificada y que la información no recibe un análisis individual sujeto a sujeto, sino que en el cuestionario, y en los métodos de encuesta en general, se opera con grandes cantidades de sujetos.

— Es necesario indicar con qué talante y actitud se debe cumplimentar el cuestionario. Por lo general, se suele pedir sinceridad en la respuesta y que el encuestado piense y reflexione antes de contestar. No está de más aclarar, especialmente si se trata de escolares, que el cuestionario no es un examen, por lo que no hay respuestas «correctas» o «incorrectas», sino respuestas que se aproximan más o menos a lo que el encuestado opina, siente o piensa.

— Dentro de la introducción, es también importante indicar la forma arbitrada para devolver la información. Bien una dirección de contacto, bien un número de teléfono o cualquier indicación que haga saber al encuestado que va a tener la posibilidad de acceder parcial o totalmente a los resultados encontrados. Si esta indicación no se hace en la introducción del cuestionario, debe al menos realizarse verbalmente o en la carta que, en ocasiones, acompaña a los cuestionarios por correo.

— Se deben ofrecer indicaciones sobre cómo se cumplimenta el cuestionario y cómo tiene que indicar su respuesta el sujeto. En el caso en que se estime necesario, se

puede poner un ejemplo de pregunta ya cumplimentada, prestando atención a que la pregunta-ejemplo no sea ninguna de las que contiene el cuestionario (para evitar que el sujeto crea que debe contestar en esa pregunta como en el ejemplo, lo cual es improbable que suceda pero puede ser el caso en algunos sujetos que no han comprendido bien las instrucciones).

— Por último, la introducción debe finalizar con una fórmula de agradecimiento por la participación.

En nuestra opinión, las preguntas biográfico-demógraficas, es decir, aquellas referidas a datos identificativos del sujeto, deben situarse al comienzo del cuestionario, tras la introducción. De esta forma, es difícil que los sujetos olviden consignar esta información y además se consigue un efecto de transición entre la actitud «pasiva» de leer la introducción e instrucciones a la actitud «activa» de contestar a las preguntas. Algunos autores (por ejemplo, GARCÍA FERRANDO, 1986), no obstante, aconsejan dejarlas para el final del instrumento, ya que así permiten concluirlo con una sensación más relajada.

Respecto a la organización y distribución de las preguntas, GARCÍA FERRANDO (1986) recomienda tener en cuenta tres aspectos: el logro de una introducción apropiada, la transición fácil y razonable de un tema a otro y la formulación de una adecuada conclusión del instrumento. Este autor plantea lo siguiente:

> «Es conveniente que el cuestionario comience con una serie de preguntas que no planteen problemas subsiguientes. A continuación, conviene que se realice una aproximación gradual al problema central. De este modo se le introduce a la persona entrevistada no sólo en el problema a investigar, sino también en el papel del sujeto que se enfrenta con un instrumento de análisis. Después vienen las preguntas más complejas o con mayor carga emocional, que conviene formular cuando la persona entrevistada se encuentra en una situación que corresponda a la naturaleza de los estímulos. Finalmente, conviene dedicar la última parte del cuestionario a preguntas que representen estímulos más fáciles y que permitan tanto al entrevistador como al entrevistado terminar la entrevista con una sensación más relajada» (GARCÍA FERRANDO, 1986:144).

De esta forma, al ser las primeras preguntas más generales se impide que las respuestas a ellas condicionen o sesguen las que se presentan con posterioridad. Esto es lo que se denomina como *secuencia de embudo* (CABRERA y ESPÍN, 1986).

La extensión final del cuestionario es un aspecto que debe tenerse muy en cuenta, no sólo en lo relativo al número de preguntas, sino también en cuanto al número de páginas que ocupa el instrumento. Dice FOX (1981:618): «cada página que se añade al cuestionario reduce la proporción de sujetos que lo completan». Esto, a veces, es una cuestión relacionada con el diseño de la página: el tipo de letra, su tamaño, la estructura y distribución de los párrafos... Han de cuidarse estas cuestiones estilísticas a fin de seleccionar aquel formato que minimiza la cantidad de folios necesario para la presentación del cuestionario, sin resultar visualmente demasiado cargado o prácticamente ilegible. Agrupar preguntas de contenido similar, como en el cuadro 6, es una forma apropiada, no sólo para ganar espacio, sino también para facilitar la lectura y cumplimentación del cuestionario.

CUADRO 6

EJEMPLO DE AGRUPACIÓN DE PREGUNTAS CON CONTENIDOS SIMILARES

➤ **¿Ha practicado alguna vez los siguientes deportes?**

- Tenis SI NO
- Fútbol SI NO
- Baloncesto SI NO
- Golf SI NO
- Otros (indicar:_____)

Respecto a la longitud del cuestionario, aunque algunos autores plantean un margen deseable en cuanto al número de preguntas, la cuestión es compleja pues esto dependerá de otros factores. Así, MUCCHIELLI (1974) plantea que la longitud óptima está entre las 15 y 30 preguntas. Otros autores recomiendan que el cuestionario no requiera más de media hora para su cumplimentación. Sin embargo, no se puede plantear este tema aisladamente de otras cuestiones. Para el sujeto requiere mayor inversión de esfuerzo mental pensar y escribir su respuesta en las preguntas abiertas que seleccionar una opción en las cerradas, además de que estas últimas se cumplimentan en menor tiempo. Por otra parte, el contenido del instrumento o tema sobre el que se realiza el diagnóstico puede revestir mayor o menor complejidad para el encuestado. Asimismo, no hay que olvidar que las preguntas incluidas en el instrumento deben recoger todos los indicadores sugeridos en el marco conceptual del mismo. Por tanto, a la hora de decidir sobre la longitud del cuestionario, hay que equilibrar todos estos factores y tomar decisiones sobre la base del sentido común. Se debe aspirar, no obstante, a que el formato final del cuestionario sea una hoja impresa por ambas caras o, a lo sumo, dos. Puede decirse que cuantos más folios añadamos a esta cantidad, menos sujetos acabarán devolviendo el cuestionario, ya que la primera impresión que les causará es que cumplimentarlo les quitará mucho tiempo.

4. GARANTÍAS CIENTÍFICAS DE LAS TÉCNICAS E INSTRUMENTOS BASADOS EN LA ENCUESTA

Suele ser costumbre entre los investigadores y profesionales del diagnóstico que construyen y usan técnicas de encuesta, preocuparse por diseñarlas y aplicarlas de forma adecuada, teniendo en cuenta muchas de las recomendaciones y consejos que aquí se han ido enumerando. Con ello esperan conseguir que sus resultados tengan las garantías oportunas de fiabilidad y validez. No obstante, esta afirmación rara vez se acompaña de estudios sistemáticos que informen de la bondad de los instrumentos utilizados. En esta situación influyen, entre otras razones, que la mayoría de las entrevistas y los cuestionarios están diseñados para «usar y tirar». Es decir, están pensados y construidos para responder a las finalidades concretas de un diagnóstico, evaluación o investigación determinados y no para ser usados en otras ocasiones, ni por su propio diseñador ni por otros profesionales.

De esta forma resulta difícil obtener el beneficio de la comprobación por el uso repetido. Como reflexión general, con todas las matizaciones que sea oportuno añadir, quizá sea hora de que elaboremos menos instrumentos, pero de mayor calidad y situemos nuestro trabajo diagnóstico e investigador en el plano más amplio de las líneas de trabajo iniciadas y compartidas con otros profesionales. De esta forma podremos mejorar de forma continua los instrumentos ya existentes, ampliando su uso tanto en lo relativo a los sujetos en los que se aplica, como en las dimensiones teórico-conceptuales que un determinado instrumento puede englobar. Un paso previo, muy simple y sencillo, consiste en adjuntar nuestros instrumentos cuando publicamos en revistas y otros medios los resultados de una investigación o un diagnóstico. Así estarán al alcance de otros profesionales que pueden estar interesados en el mismo tema. Si a ello añadimos una referencia a las garantías que el instrumento ha demostrado reunir en cuanto a fiabilidad y validez, nuestra contribución al conocimiento científico y al avance del saber será exponencialmente mayor.

En este apartado, con el que concluye el capítulo dedicado a las técnicas e instrumentos de encuesta, queremos abordar, aunque sea de forma breve, el tema de las garantías científicas de estos procedimientos, al igual que hicimos con la observación. No obstante, dado que los criterios de calidad científica no son diferentes según el procedimiento que se use, mucho de lo expuesto en el capítulo de la observación es también aplicable a los métodos de encuesta, por lo que aquí nos limitaremos a ahondar en algunas de las aplicaciones de estos criterios a los cuestionarios y entrevistas.

Dadas las características propias de las tres técnicas abordadas en este capítulo (cuestionarios, entrevistas estructuradas y entrevistas abiertas), consideramos más adecuado tratar separadamente las cuestiones relativas a las garantías científicas de cada una de estas técnicas.

4.1. Cuestionarios

En apartados anteriores se han abordado de forma parcial las posibles fuentes de error que afectan a los cuestionarios. A todo lo que ya se ha apuntado debemos añadir un conjunto de factores que pueden incidir en las propiedades psicométricas de los cuestionarios, haciendo que sea menor su validez y su fiabilidad. JENSEN y HAYNES (1986[7]; cit. por MATUD, 1993) revisan las *fuentes de error de los cuestionarios*, distinguiendo entre las siete siguientes (siendo las tres primeras voluntarias y el resto involuntarias):

1) La deseabilidad social, que implica que las respuestas a las preguntas formuladas pueden estar influidas por los valores sociales, de forma que el sujeto tiende a presentarse a sí mismo de forma favorable.

2) Las características de la población evaluada también inciden en la validez de la información recogida a través de un cuestionario. Así, resulta necesario investigar en profundidad la diferencia en las respuestas de determinados grupos (por ejemplo, alcohólicos, sociópatas ...).

3) Las características de la demanda de la situación de evaluación, de tal forma que la exactitud de la respuesta variará en función de ella (selección para un trabajo, peritaje judicial ...).

4) La percepción errónea o incompleta de los sujetos (dificultad para recordar, tendencia a fijarse en información no relevante ...).

5) Las influencias reactivas o posibles cambios que el procedimiento de evaluación introduce en la conducta o características objeto de la misma.

6) Los sesgos en las respuestas, que pueden provenir de factores como la dependencia serial de los ítems (la respuesta a un ítem puede influir en la respuesta al siguiente) o de las tendencias de respuesta (hacia las puntuaciones inferiores, centrales o superiores).

7) Los sesgos que provienen de otras variables implicadas en la construcción y aplicación de los cuestionarios (instrucciones, lenguaje utilizado, lugar y condiciones de administración, ...).

Todos estos factores deben ser controlados a fin de garantizar que los cuestionarios que usamos en los procesos diagnósticos estén libre, en la medida de lo posible, de su contaminación. En caso contrario, la información recogida a través de ellos estará distorsionada y no nos proporcionará una base adecuada para la toma de decisiones.

Por otra parte, los cuestionarios comparten muchos aspectos en común con los tests y las pruebas objetivas. En la medida en que un cuestionario está muy estructurado y sus

preguntas y modelos de respuesta están previamente definidos y permiten la cuantificación de los fenómenos, su *fiabilidad y validez* puede analizarse por los mismos procedimientos establecidos para los tests y pruebas objetivas. A fin de no repetir conceptos, recomendamos al lector la lectura del apartado correspondiente en el capítulo 4, que será de aplicación en estos casos.

4.2. Entrevistas estructuradas

Resulta especialmente difícil valorar la calidad de la entrevista como estrategia de recogida de información. Como sugiere SILVA (1992), si la calidad de un instrumento se juzga esencialmente según sus objetivos, valorar la entrevista reviste una gran complejidad, dada la diversidad y heterogeneidad de fines y objetivos para los que se utiliza esta estrategia. GARCÍA MARCOS (1983) también plantea que la multiplicidad de funciones a las que sirve la entrevista en los procesos diagnósticos, junto con el amplio número de variables que inciden en ella (variables relativas al examinado, al examinador, al tema, etc.), han dificultado un acercamiento científico a la calidad de las entrevistas.

A ello deberíamos añadir que, como se ha indicado, las entrevistas pueden presentar diversos grados de estructuración. Así, en la medida en que la entrevista está muy estructurada, los procedimientos para su validación no tienen por qué ser diferentes a los usados en los cuestionarios, si bien para ello sería necesario hacer una aplicación masiva del instrumento de entrevista diseñado, ya que dichos procedimientos se basan en análisis estadísticos que requieren trabajar con datos procedentes de muchos sujetos. Esto suele ser poco habitual, ya que para trabajar con muestras amplias, el cuestionario resulta una estrategia menos costosa y más efectiva.

En el cuadro 7 hemos sintetizado la opinión de diversos autores sobre los posibles sesgos o *fuentes de sesgo* que pueden afectar a la calidad de las entrevistas. En general, puede observarse que las fuentes de sesgo provienen bien del entrevistador, bien del entrevistado, o bien del método de entrevista en sí mismo o de la situación concreta en que ésta tiene lugar.

CUADRO 7

RESUMEN DE LAS POSIBLES FUENTES DE SESGO EN LAS ENTREVISTAS

FUENTES DE SESGO EN LA ENTREVISTA		
García Marcos (1983)	**Ávila Espada (1990)**	**Mateu y Fuentes (1992)**
— Del entrevistador — Del entrevistado — Del método	— Fenómenos y efectos del entrevistador — Fenómenos y efectos del entrevistado — Fenómenos y efectos de la situación de entrevista	— Habilidad del entrevistador — Estatus de entrevistado y entrevistador — Efectos de reactividad — Contenidos de la entrevista — Desconocimiento o falta de información

Para reducir, en la medida de lo posible, la incidencia de estos sesgos es necesario tener en cuenta todas las recomendaciones que hemos realizado en apartados anteriores, ya que todas ellas se orientan a mejorar la calidad de las entrevistas, e influyen en su fiabilidad y validez. En el cuadro 8 recogemos un conjunto de recomendaciones aportadas por SILVA (1992) tras revisar las características que han demostrado mejorar la calidad de la información recogida a través de las entrevistas. Se puede observar que, básicamente, se refieren a cuestiones ya planteadas en este capítulo.

CUADRO 8

RECOMENDACIONES DE SILVA (1992) PARA MEJORAR LA OBJETIVIDAD, FIABILIDAD Y VALIDEZ DE LAS ENTREVISTAS

RECOMENDACIONES PARA MEJORAR LA CALIDAD DE LAS ENTREVISTAS

1) Para asegurar una mayor objetividad, conviene...

- — Estructurar preguntas y respuestas
- — Usar categorías de registro claramente definidas
- — Ordenar las respuestas en esquemas o pautas
- — Que exista unidad de perspectivas en el juicio de distintos entrevistadores
- — Entrenar en el registro y tabulación de la información

2) Para asegurar una mayor fiabilidad, conviene...

- — Estimar las bases motivacionales de la respuesta
- — Utilizar racimos de preguntas en lugar de preguntas aisladas
- — La clara especificación temporal de las preguntas
- — Preguntar por acontecimientos objetivos o indicadores observables de la conducta
- — Facilitar al entrevistado la posibilidad de contrastar la adecuación de sus respuestas

3) Para asegurar una mayor validez, conviene...

- — Dar instrucciones y aclaraciones que sitúen al entrevistado en los diversos temas
- — Que las preguntas sean transparentes y el lenguaje comprensible
- — Facilitar diversas alternativas de respuesta
- — Usar preguntas que se circunscriban a intervalos temporales o a acontecimientos determinados
- — Dar preferencia a los indicadores observables de la conducta
- — Intentar detectar y neutralizar tendencias de respuesta
- — Asegurar la confidencialidad de la información y el secreto profesional
- — Abrir la posibilidad de contrastar la respuesta mediante otras fuentes de información
- — Evitar sugerencias

En cuanto a la *fiabilidad de las entrevistas*, se ha estudiado preferentemente la fiabilidad interjueces y la fiabilidad test-retest. La fiabilidad interjueces se entiende como el grado de acuerdo entre las estimaciones y valoraciones efectuadas por distintos observadores para una misma entrevista. En algunas ocasiones se ha utilizado este mismo concepto para referir el acuerdo existente entre dos entrevistas hechas por separado a un mismo sujeto por parte de dos entrevistadores distintos (PEÑATE, 1993). Esta concordancia entre evaluadores, también llamada «fiabilidad del evaluador» (SILVA, 1992), se ha estudiado preferentemente en el diagnóstico psicopatológico y en la

evaluación de las características de la personalidad, por lo que cabe cuestionar la aplicación de los resultados en otros contextos, como es el caso del campo socio-educativo. Se encuentra que el acuerdo entre evaluadores es tanto más alto cuanto menos específico es el diagnóstico. Al mismo tiempo, los coeficientes son más altos cuando se valoran aspectos conductuales, objetivos y accesibles a la observación, obteniéndose coeficientes más bajos para las valoraciones de impresiones subjetivas e interpretaciones personales (MATEU y FUENTES, 1992; SILVA, 1992). El grado de entrenamiento y la experiencia de los entrevistadores en la utilización de sistemas de categorías bien delimitados y definidos también son factores que inciden muy positivamente en la fiabilidad inter/intrajueces (ÁVILA ESPADA, 1990).

A pesar de todo ello, GARAIGORDOBIL (1998) y PEÑATE (1993) sugieren que el acuerdo intercalificadores no debe ser tomado como un criterio de bondad psicométrica más, sino como un requisito indispensable para que una entrevista pueda llevarse a cabo con garantías.

Según SILVA (1992), en la investigación sobre la fiabilidad test-retest o concordancia entre entrevistas realizadas en momentos diferentes, también se encuentran coeficientes más altos cuando la información es más objetiva. La duración del intervalo entre el test y el retest es otra cuestión que afecta a la concordancia, cumpliéndose la tendencia psicométrica general según la cual, en intervalos cortos, los resultados tienden a ser más satisfactorios. Por su parte, PEÑATE (1993) señala que razones como el muestreo insuficiente de contenidos, el tamaño de la muestra, o la propia naturaleza cambiante del objeto evaluado se han utilizado para justificar los bajos niveles de fiabilidad test-retest alcanzados por la investigación.

GARCÍA MARCOS (1983) apunta otra forma de estimar la fiabilidad de las entrevistas; en este caso se trata de comprobar la consistencia o congruencia interna de la información suministrada por el mismo entrevistado en varias partes de la entrevista. Según este autor, este tipo de fiabilidad apenas ha despertado el interés de la investigación.

En relación con la *validez de las entrevistas*, se ha considerado que ésta se refiere al grado en que el informe verbal del sujeto o sujetos durante la entrevista refleja adecuadamente su conducta en el ambiente natural (HAYNES, 1978[8]; cit. por GARAIGORDOBIL, 1998). Esta es la razón por la que se ha estudiado preferentemente el valor predictivo o criterial de la información obtenida en la entrevista (validez criterial).

Sin embargo, también se ha considerado necesario estudiar la validez de contenido, en especial en relación con el carácter más o menos completo de la información recogida a través de formatos semiestructurados o estructurados de entrevista. No obstante, es difícil estimar este aspecto en las entrevistas más abiertas, que poseen menor nivel de estructuración. Las entrevistas estructuradas —y las semiestructuradas que tienden más al polo de la estructuración— deben explicitar los contenidos que pretender evaluar y la relación de preguntas a través de las cuales pretenden evaluarlo. Sólo así resulta posible

analizar las evidencias de la validez de contenido del instrumento finalmente diseñado, con procedimientos similares a los ya apuntados para la observación (por ejemplo, el juicio de expertos).

En lo que se refiere a la validez criterial, en algunos casos se ha empleado el criterio de acuerdo intercalificadores como indicio de esta faceta de la validez (GARAIGORDOBIL, 1998; PEÑATE, 1993). No obstante este procedimiento ha sido desestimado por algunos investigadores (ENDICOTT, 1981[9]; cit. por GARAIGORDOBIL, 1998), que alegan que se trata de un procedimiento para estimar la fiabilidad en lugar de la validez. En el campo de la psicología clínica se ha utilizado otro procedimiento para la estimación de la validez criterial consistente en evaluar en qué medida una entrevista puede reflejar la patología de alguien previamente diagnosticado. No obstante, según informan GARAIGORDOBIL (1998) y SILVA (1992), la heterogeneidad que reina en la información referente a la validez criterial de los datos de entrevista es enorme, por lo que es difícil encontrar alguna tendencia general. Según estos mismos autores, los coeficientes mejoran a medida que la entrevista se acerca a criterios más objetivos, así como a sistemas diagnósticos respaldados a su vez por una mayor validez y mejor estructurados. No obstante, se constatan distorsiones según el tipo de información requerida en la entrevista (según ésta sea más o menos íntima o comprometida), así como según el tipo de finalidad de la entrevista (y lo que se «juega» el sujeto al cual se aplica).

Como se puede observar, las escasas aportaciones en torno a la validez de las entrevistas provienen del campo de la psicología clínica, ya que es evidente que la entrevista es un procedimiento ampliamente usado en la evaluación e intervención psicológica. Dentro de este campo, ÁVILA ESPADA (1990) realiza una revisión en torno a esta cuestión, concluyendo que los distintos tipos de validez de la entrevista dependen estrechamente de los objetivos de la misma, del contexto en que se realiza y del sistema teórico o marco referencial del entrevistador. En el campo de la evaluación laboral, MCDANIEL et al. (1994) también han realizado meta-análisis de estudios que vienen a poner de manifiesto que la validez en la entrevista depende del contenido de la misma, de su grado de estructuración y de la naturaleza del criterio que se utiliza.

Resumiendo lo indicado en torno a las características psicométricas de la entrevista, es necesario valorar positivamente el aumento del número de investigaciones en torno a esta cuestión. En palabras de PEÑATE (1993:214), ello «ha permitido superar la idea de la entrevista como la "gran desconocida" de los instrumentos diagnósticos». De cualquier manera, al igual que ocurre con los tests, la observación y con otros procedimientos, la bondad psicométrica de una técnica de diagnóstico y evaluación no es extrapolable a todos los instrumentos concretos que pertenecen a la misma categoría, sino que cada uno de ellos debe someterse a estudio.

4.3. Entrevistas abiertas

La aplicación de los criterios habituales de calidad científica —fiabilidad y validez— a las entrevistas abiertas o en profundidad es una cuestión aún más compleja. Como plantea ALONSO (1994:229), este tipo de entrevistas «es refractaria a cualquier criterio cientificista de definición de la herramienta metodológica». En ella, y siguiendo con el mismo autor:

— No existe regla fija ninguna sobre la forma de realizar la entrevista ni la conducta del entrevistador.

— Toda entrevista es producto de un proceso interlocutorio que no se puede reducir a una contrastación de hipótesis y al criterio de falsación.

— Los resultados de la entrevista no tienen posibilidad de generalización indiscriminada ni mucho menos de universalización.

Por esta razones, la entrevista abierta sólo se puede juzgar, como otras técnicas cualitativas, por sus resultados finales, por la riqueza heurística de las producciones discursivas obtenidas mediante ella y, en general, por los mismos criterios de credibilidad, transferibilidad, dependencia y confirmabilidad (GUBA, 1983; LINCOLN y GUBA, 1985), analizados en el capítulo anterior.

Por tanto, la validez y fiabilidad de las entrevistas abiertas es una cuestión que plantea serias dificultades. Es interesante citar aquí, la aportación de KITWOOD (1988[10]; cit. por COHEN y MANION, 1990:391):

> «En la misma proporción que aumenta la "fiabilidad" mediante la racionalización disminuiría la "validez". Puesto que el principal propósito de usar una entrevista en investigación es que se cree que, en un encuentro interpersonal, la gente está más proclive a descubrir aspectos de sí misma, sus pensamientos, sus sentimientos y valores, que en una situación menos humana. Al menos, para ciertos propósitos es necesario generar una clase de conversación en la que el "informante" se sienta cómodo. En otras palabras, el distintivo del elemento humano es necesario en la entrevista para su "validez". Cuanto más racional, calculador y alejado se haga el entrevistador, menos probable es que se perciba la entrevista como un intercambio amistoso y también es más probable que sea más calculada la respuesta».

WOODS (1989) ofrece numerosas indicaciones para lo que él denomina como «la validación de los relatos» (p. 98), refiriéndose con ello a la necesidad de conocer si el informante dice la verdad. Entre este conjunto de indicaciones, resaltamos las siguientes:

— Procurarse varias versiones del relato, tanto con el informante como con respecto a terceras personas. Esto implica volver a entrevistar al informante en diferentes ocasiones (así vemos si su opinión varía en situaciones y momentos distintos) para obtener una muestra completa de sus opiniones. Asimismo, implica entrevistar a otras personas a fin de cotejar distintas versiones del relato. En este respecto son de gran interés los informantes clave (aquellas personas que ayudan al investigador identificando e interpretando muchos elementos del ámbito que éste estudia).

— Procurar, si resulta aconsejable, la «validación del demandado», esto es, la devolución del relato al informante para su apreciación.

— Procurar la triangulación, siendo éste el medio más importante, en opinión de WOODS (1989), para la validación. Proporciona puntos de vista alternativos y

complementarios. Para WOODS (1989), el vínculo más fuerte se produce cuando las entrevistas van acompañadas de observación, ya que ambos métodos combinados permiten una descripción más exhaustiva, así como detectar distorsiones y exageraciones.

Algunas de estas ideas están implicadas en el concepto de «controles cruzados» propuesto por TAYLOR y BOGDAN (1990:127). A través de este concepto se refieren a comparar las informaciones aportadas por el sujeto con los datos provenientes, bien de otros informantes, bien de algún tipo de documento escrito.

Por último, en este tipo de entrevistas, hay que prestar especial atención a lo que se ha denominado como *efecto greenspoon* (PEÑATE, 1993), que ha sido estudiado en el ámbito de las entrevistas psicológicas clínicas. Así, las producciones lingüísticas del entrevistado son susceptibles de modificarse en su tasa de ocurrencia a través de la aprobación verbal del interlocutor. De esta forma, las típicas expresiones que habitualmente utilizamos para demostrar nuestra atención y/o aprobación («hummm, ya, sí ...») pueden condicionar lo que el entrevistado dice. Esto lleva a cuestionar si en cualquier proceso de entrevista no se está de alguna forma condicionando al entrevistado a través de las muestras de *feed-back* que ofrece el entrevistador.

En definitiva, pese a la dificultad para aplicar en las entrevistas abiertas y, en general, en las técnicas cualitativas para el diagnóstico, los criterios habituales de fiabilidad y validez, parece claro que existen procedimientos que contribuyen a afianzar y mejorar la calidad de la información aportada por estas técnicas. La utilización de estos procedimientos por parte del profesional del diagnóstico es necesaria, ya que utilizar estrategias cualitativas en el diagnóstico educativo no consiste en aplicar la regla de «todo vale», en cuyo caso restaríamos cientificidad a los procesos desarrollados.

Referencias bibliográficas

ALONSO, L. E. (1994). Sujeto y discurso: el lugar de la entrevista abierta en las prácticas de la sociología cualitativa. En J.M. Delgado y J. Guitérrez (coords.), *Métodos y técnicas cualitativas de investigación en ciencias sociales* (pp. 225-240). Madrid: Síntesis.

ÁVILA ESPADA, A. (1990). La entrevista clínica. En A. Avila (ed.), *Evaluación psicológica clínica*. Madrid: Facultad de Psicología de la Universidad Complutense.

BUENDÍA, L. (1994). Técnicas e instrumentos de recogida de datos. En M.P. Colás y L. Buendía, *Investigación educativa* (pp. 201-248). Sevilla: Alfar.

CABRERA, F. A. y ESPÍN, J. V. (1986). *Medición y evaluación educativa*. Barcelona: PPU.

COHEN, L. y MANION, L. (1990). *Métodos de investigación educativa*. Madrid: La Muralla.

FERNÁNDEZ BALLESTEROS, R. (1992). Los autoinformes. En R. Fernández Ballesteros (ed.), *Introducción a la evaluación psicológica* (pp. 218-251). Madrid: Pirámide.

FONTANA, A. y FREY, J. H. (1988). Interviewing: The art of science. En N.K. Denzin y Y.S. Lincoln (ed.), *Handbook of qualitative research*. Beverly Hills, CA: Sage.

FOX, D. J. (1981). *El proceso de investigación en educación*. Pamplona: EUNSA.

GARCÍA FERRANDO, M. (1986). La encuesta. En M. García Ferrando, J. Ibáñez y F. Alvira (comp.), *El análisis de la realidad social. Métodos y técnicas de investigación* (pp. 123-152). Madrid: Alianza Editorial.

GARCÍA JIMÉNEZ, E. (1997). *La recogida de información*. Material de la asignatura Diagnóstico Pedagógico, inédito. Sevilla: Facultad de Ciencias de la Educación.

GARCÍA MARCOS, J. A. (1983). La entrevista. En R. Fernández Ballesteros (coord.), *Psicodiagnóstico*, *volumen I* (pp. 355-395). Madrid: UNED.

GARAIGORDOBIL, M. (1998). *Evaluación psicológica. Bases teórico-metodológicas, situación actual y directrices de futuro*. Salamanca: Amarú Ediciones.

GUBA, E. G. (1983). Criterios de credibilidad en la investigación naturalista. En J. Gimeno y A. Pérez (ed.), *La enseñanza: su teoría y su práctica* (pp. 148-165). Madrid: Akal.

HAGUE, P. N. y JACKSON, P. (1994). *Cómo hacer investigación de mercado*. Bilbao: Deusto.

KORNHAUSER, A. y SHEATSLEY, P. B. (1976). Construcción de cuestionarios y procedimientos de entrevista. En C. Selltiz, M. Jahoda, M. Deutsch y S.W. Cook, *Métodos de investigación en las relaciones sociales* (pp. 601-646). Madrid: Rialp.

LINCOLN, Y. y GUBA, E. (1985). *Naturalistic enquiry*. Beverly Hills, CA: Sage.

MANZANO, V. G., ROJAS, A. J. y FERNÁNDEZ, J. S. (1996). *Manual para encuestadores*. Barcelona: Ariel.

MATEU, C. y FUENTES, I. (1992). *La entrevista de evaluación*. Valencia: Nau llibres.

MATUD, P. (1993). Cuestionarios e inventarios de autoinforme. En W. Peñate, P. Mataud e I. Ibáñez, *Evaluación psicológica: Concepto y técnicas de análisis* (pp. 283-210). Valencia: Promolibro.

McDANIEL, M. A. et al. (1994). The validity for employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology, 79* (4), 599-616.

MERRIAM, S. B. (1988). *Case study research in education: a qualitative approach*. San Francisco, CA: Jossey-Bass.

MORGAN, H. H. y COGGER, J. W. (1989). *Manual del entrevistador*. Madrid: Tea.

MUCCHIELLI, R. (1974). *El cuestionario en la encuesta psico-social* (4ª edición). Madrid: Ibérico Europea de Ediciones.

ORTÍ, A. (1986). La apertura y el enfoque cualitativo o estructural: la entrevista abierta semidirectiva y la discusión de grupo. En M. García Ferrando, J. Ibáñez y F. Alvira (comp.), *El análisis de la realidad social. Métodos y técnicas de investigación* (pp. 153-185). Madrid: Alianza Editorial.

PATTON, M. (1990). *Qualitative evaluation and research methods*. Newsbury Park, CA: Sage Publications.

PEÑATE, W. (1993). La entrevista. En W. Peñate, P. Mataud e I. Ibáñez, *Evaluación psicológica: Concepto y técnicas de análisis* (pp. 173-222). Valencia: Promolibro.

RODRÍGUEZ, J. M. y BLANCO, A. (1986). Técnicas psicométricas. En A. Blanco (ed.), *Apuntes de psicodiagnóstico*. Valencia: Promolibro.

RODRÍGUEZ SUTIL, J. C. (1994). La entrevista psicológica. En J.M. Delgado y J. Guitérrez (coord.), *Métodos y técnicas cualitativas de investigación en ciencias sociales* (pp. 241-255). Madrid: Síntesis.

ROSSELL, T. (1989). *La entrevista en el trabajo social*. Barcelona: Escuela Universitaria de Trabajo Social, Hogar del libro.

SELLTIZ, C., JAHODA, M., DEUTSCH, M. y COOK, S. W. (1976). *Métodos de investigación en las relaciones sociales*. Madrid: Rialp.

SILVA, F. (1992). La entrevista. En R. Fernández Ballesteros (coord.), *Introducción a la evaluación psicológica*, v*olumen I,* (pp. 252-278). Madrid: Pirámide.

SPRADLEY, J. (1979). *The etnographic interview*. New York: Holt, Rinehart & Winston.

TAYLOR, S. J. y BOGDAN, R. (1990). *Introducción a los métodos cualitativos de investigación*. Buenos Aires: Paidós.

WOODS, P. (1989). *La escuela por dentro. La etnografía en la investigación educativa*. Madrid: Paidós/MEC.

CAPÍTULO

4

INTRODUCCIÓN AL DISEÑO Y APLICACIÓN DE TESTS Y PRUEBAS OBJETIVAS

1. LA MEDIDA EN EL DIAGNÓSTICO Y LA EVALUACIÓN

STEVENS (1951) proporciona una definición ya clásica del acto de medir como la asignación de numerales a objetos o hechos de acuerdo con reglas o normas fijas. Ahora bien, ante esta definición debemos hacer dos matizaciones:

— Por un lado, cuando hablamos de «asignar números», en realidad nos referimos a asignar las *propiedades de los números* (CABRERA y ESPÍN, 1986). Estas propiedades pueden ser: igualdad o desigualdad, orden, igualdad de diferencia e igualdad de razón. Esto nos recuerda los distintos niveles de medida con los que trabajamos en educación y otras disciplinas (nominal, ordinal, de intervalo y de razón). La asignación de estas propiedades de los números a lo que medimos es lo que nos permite conocer si dos sujetos poseen o no la misma «cantidad» de un rasgo, si son diferentes o iguales en una propiedad, etc.

— Por otro, al medir no asignamos números a los objetos o eventos, sino a los atributos o propiedades de esos objetos o eventos (CAMACHO y SÁNCHEZ, 1997). Así, nunca medimos a las personas, sino sus características o propiedades (MEHRENS y LEHMANN, 1982), como su altura, peso, conocimientos, actitudes, etc.

Sin embargo, para que pueda llevarse a cabo una medición, debe existir un *isomorfismo* entre los objetos o hechos que se pretenden comparar y el sistema numérico que utilizamos para ello (GONZÁLEZ BLASCO, 1986). Dicho de otra forma, la regla que utilizamos para asignar números a las propiedades de los objetos o características de las personas, debe garantizar que las relaciones entre los números mantengan las relaciones que existen entre los valores de los atributos que estamos midiendo.

Según qué normas se sigan en la asignación de números a los atributos o propiedades se pueden establecer las distintas escalas de medida, que son: *nominal* (sólo permite realizar operaciones de igualdad o diferencia en el atributo); *ordinal* (además de la anterior, permite conocer si dos sujetos u objetos poseen en mayor o menor medida el atributo); de *intervalo* (en la que, a las comparaciones anteriores, podemos añadir la posibilidad de comparar diferencias entre valores del atributo); y, finalmente, la de *razón* (en la que se pueden establecer todas las comparaciones anteriores pero además permite comparar proporciones entre pares de valores del atributo).

Por lo tanto, la definición de STEVENS (1951) permite incluir en el concepto de medida las variables nominales y es, por tanto, más fácilmente aplicable a las variables educativas. Ahora bien, como plantea PÉREZ JUSTE (1991:11), se puede considerar que la «medición admite medición», en tanto puede haber mediciones más exigentes que otras, y distintas clases de medidas. Podemos entender, bajo este supuesto, que existe una medición en sentido amplio y una medición en sentido estricto.

Hablamos de medición en sentido amplio siempre que, según la definición de STEVENS (1951), asignamos unos números y sus propiedades a los atributos y características de los objetos y sujetos que estamos evaluando —bajo el supuesto de isomorfismo que hemos apuntado—. La cuantificación es, por tanto, el concepto de medición defendido por Stevens y que nosotros consideramos como medida en sentido amplio.

Sin embargo, algunas de las medidas proporcionadas bajo este concepto amplio de medición son muy burdas y difícilmente comparables entre sí. Por ejemplo, si «medimos» (en sentido amplio) el sexo de una persona y asignamos un 1 a todos los hombres y un 2 a todas las mujeres, nos encontramos con un ejemplo clásico de variable medida en escala nominal. En este caso, al hablarse sólo de igualdad o diferencia en la posesión del atributo (hombre o mujer), la medida que obtenemos es demasiado genérica como para que pueda producir alguna significación. No podríamos, por poner un ejemplo, decir que puesto que «mujer» equivale a 2, una mujer es el doble que un hombre, ni que un hombre tiene la mitad del atributo «sexo» que una mujer.

Cuando nos referimos a las variables ordinales nos sucede un poco de lo mismo, aunque el tema es más complejo. Si preguntamos a alguien si recicla su basura y le damos a escoger entre las opciones «nada», «algo», «bastante» y «mucho», podemos inferir que si escoge la opción «mucho» recicla más basura que el que ha elegido «nada». Si cuantificamos su respuesta en 1, 2, 3 y 4, según conteste «nada», «algo», «bastante» o «mucho», podemos decir que aquel sujeto con 4 recicla más basura que el sujeto que ha obtenido un 2. Ahora bien, ¿cuánta basura «más» recicla el primero respecto del segundo? Podría suceder que el que ha contestado 4 recicle dos kilos de papel al año y por eso considera que recicla «mucho». Sin embargo, nada se opone a que el sujeto que tiene un 2 recicle cinco kilos de papel al año, pero puede ser más exigente con el reciclaje de basura y esa cantidad le parece «algo» y no «mucho», ni «bastante». Es por ello que el nivel de medida ordinal sigue siendo una estimación demasiado «gruesa» de la cantidad de un atributo.

Ahora bien, cuando nosotros presentamos a dos sujetos distintos un conjunto de tareas para ver cuál de los dos consigue resolver adecuadamente más cantidad de ellas, la cuestión es diferente. Si asignamos a cada sujeto un punto cada vez que resuelve bien una tarea, el que uno obtenga 10 puntos y otro 8 sí que implica que el primero ha resuelto adecuadamente dos tareas más. En este caso, disponemos de una *unidad de medida*, de forma que el número (8 o 10) está indicando la cuantía de la diferencia, y esto es, en nuestra opinión, lo propio y específico de la medida en sentido estricto, frente a la cuantificación o medida en sentido amplio. Es decir, desde nuestro punto de vista, la medida en sentido estricto se consigue cuando somos capaces de cuantificar las diferencias entre los valores de atributo porque disponemos de una unidad (intervalos iguales entre puntuaciones contiguas); así, entre dos puntuaciones cualesquiera, una diferencia observada en números equivale a una diferencia idéntica en el rasgo medido, y

esto en cualquier zona de la escala en que se encuentre la diferencia. Esto sólo es posible cuando la escala de medida de una variable es de intervalo o razón.

Si bien las variables educativas rara vez alcanzan el nivel de medición de razón (¿cómo detectamos dónde está el cero absoluto en una actitud, o en la inteligencia, o en los intereses profesionales...?), el nivel de intervalo sí resulta factible, aunque complejo, de alcanzar en el diagnóstico educativo[11]. Con esto nos adentramos en una cuestión inseparable de la conceptualización de la medida, que se refiere a los procedimientos y condiciones necesarios para llegar a la medición de intervalo.

Alcanzar un nivel de medida de intervalo es una cuestión difícil cuando los atributos que queremos medir no son directamente observables, como sucede en el campo de la educación. Si tomamos como ejemplo las actitudes ante el medio ambiente, ¿cómo podemos saber el grado exacto en que una persona tiene «más actitudes» hacia el medio ambiente que otra? Las actitudes no se observan, por tanto, ¿cómo podemos pretender medirlas? En esta línea, CAMACHO y SÁNCHEZ (1997) afirman que el objetivo que perseguimos a la hora de medir es elaborar indicadores de los conceptos teóricos objeto de medición que estén asentados en el mundo empírico.

Las actitudes constituyen un constructo o, lo que es lo mismo, un concepto teórico que ha desarrollado la comunidad científica para explicar los fenómenos que observa. Así, si estudiamos el comportamiento que las personas tienen respecto al medio ambiente, observaremos que algunas de ellas tiran papeles al suelo y otras no; observaremos también que algunas separan los diferentes tipos de residuos y otras no; etc. Para explicar por qué unas personas se comportan de un modo frente al medio ambiente y otras de forma distinta, los teóricos han desarrollado el concepto de *actitud* ante el medio ambiente. Según una persona tenga más o menos actitud ante este tema, orientará su comportamiento de una forma u otra (hacia el reciclado o el no reciclado, hacia la limpieza del entorno o la no limpieza del mismo, etc.). Por lo tanto, lo que se ha hecho es observar unos comportamientos relacionados entre sí e inferir, a partir de ellos, un rasgo, característica o cualidad —la actitud—. Como la actitud no es directamente observable, podemos recurrir a los comportamientos que sí lo son y que se suponen relacionados con ella: separar residuos, usar las papeleras, economizar el agua... En la medida en que una persona manifieste o no estos comportamientos podemos inferir que tiene la actitud, en este caso, ante el medio ambiente.

Ahora bien, ¿cuántos comportamientos y cuáles hay que manifestar para tener más o menos actitud ante el medio ambiente? Aquí es donde radica la verdadera complejidad y dificultad de conseguir medidas en sentido estricto en el campo de la educación. Para dar una respuesta más o menos satisfactoria, deberíamos hacer un listado con todas aquellas conductas que son indicativas de una actitud positiva ante el medio ambiente. En nuestro ejemplo hemos proporcionado algunas (tirar papeles a la papelera, separar tipos de residuo...). Sin embargo, muchas otras conductas serían también relevantes (no producir ruidos molestos, cerrar un grifo abierto, reprender a quien tira papeles al suelo, gastar

menos cantidad de papel, etc.). Y no sólo eso; habría que definir muy bien cada una de estas conductas porque si no, podemos darlas como indicativas de que existe una actitud positiva ante el medio ambiente, cuando no es así. Por ejemplo, siempre que cerramos un grifo, no lo hacemos para cuidar el medio ambiente. Siempre que estamos en silencio, sin producir ruidos desagradables, tampoco. Por otra parte, a veces nos vemos obligados a producir ruidos desagradables. Si no fuera así, un obrero que taladra el asfalto para construir una cañería de gas tendría unas actitudes medioambientales muy negativas.

Esto nos lleva a pensar que la medición en sentido estricto requiere el establecimiento y construcción de una regla. Es igual que si queremos medir cualquier atributo observable. Necesitaríamos inventar un sistema de medida que satisfaga a los distintos observadores, para que así pudiera estimarse la «cantidad» de atributo. El problema con que nos encontramos en las ciencias humanas y sociales es que esta tarea de *construir una regla* es una cuestión teórica y conceptual. PÉREZ JUSTE (1991:16) afirma que «la elaboración de la regla exige, para que ésta sea correcta, una teoría sobre lo que ha de medirse y, a ser posible, la contrastación de tal teoría con la realidad para poner de manifiesto su plausibilidad».

En el ejemplo sobre las actitudes ante el medio ambiente que hemos expuesto, hemos visto que es necesario identificar, seleccionar y definir una serie de conductas observables para poder decir que se mide el constructo. Esta selección, definición y acotación de indicadores (conductas observables) debe realizarse con sumo cuidado y atención a los más mínimos detalles porque, en caso contrario, dudaríamos de la *regla* que se ha utilizado para medir y, como consecuencia, dudaríamos de la medida finalmente obtenida. Recordemos el caso del obrero y la taladradora: nuestra regla debe ser tal que discrimine estas situaciones y ofrezca una medida de las actitudes ante el medio ambiente y no de otras cosas irrelevantes o no relacionadas con ellas.

En función de esta argumentación, parece claro que la medición en sentido estricto no es algo que provenga sólo de la cuantificación, sino que tiene lugar como consecuencia de unos procesos teóricos que permiten la construcción de una regla, más o menos reconocida y aceptada. Sólo en la medida en que seamos capaces de crear una regla de correspondencia unívoca entre algunas conductas y las valoraciones numéricas de las mismas, podremos hablar de medición en sentido estricto.

Por lo tanto, lo que permitirá la medición no es en sí misma una escala de intervalos (o de razón) *per se*, sino un sistema teórico que da sentido y significado a esa escala. Ese sistema teórico es el que se requiere para medir, y sólo los instrumentos desarrollados escrupulosamente desde un sistema teórico tal serán susceptibles de aportar una medida, y no una mera cuantificación.

Bajo esta afirmación, las técnicas e instrumentos que hemos presentado en los capítulos anteriores se aproximan más o menos a nuestro concepto de medición (a partir de ahora, hablaremos siempre de medición en sentido estricto). En dichos capítulos hemos presentado instrumentos como el diario o la entrevista abierta que ni siquiera

permiten la cuantificación (o, aun cuando la permiten, cuantificar significaría perder la riqueza de la información recabada). Otros procedimientos (los sistemas categoriales de observación, la entrevista estructurada y el cuestionario) se aproximan bastante al concepto de medición. En ellos, hemos insistido en la necesidad de un marco teórico que los oriente y dé sentido, y que se concrete en un conjunto de indicadores observables de las dimensiones a evaluar.

A pesar de ello, en nuestra opinión, ni los cuestionarios ni muchos sistemas categoriales permiten la medición. Dejando de lado el hecho de que, en la mayor parte de los casos, sus constructores no siempre son tan rigurosos como para seguir el proceso descrito para su elaboración, aun cuando este proceso se sigue, el grado de rigurosidad y detalle con que se hace no es suficiente para asegurar la medición. Tomando como ejemplo el fragmento de cuestionario sobre hábitos lectores que se presenta en el anexo 4, pese a que las preguntas son indicadores observables de las «preferencias literarias» de los adolescentes, ¿cómo podemos estar seguros de que se han elegido los mejores indicadores para esta dimensión?, ¿acaso no se podrían haber formulado otros indicadores?, ¿hasta qué punto los indicadores formulados son homogéneos entre sí para poder «sumar» sus efectos y otorgar, en función de esta suma, una puntuación diferencial a cada sujeto según su respuesta?... Esta última pregunta es de gran importancia. Siguiendo con el ejemplo, ¿qué demuestra «más» cantidad de preferencias literarias: leer muchos tipos de formatos impresos (libros, diarios, revistas...) o leer muchas veces al año uno de ellos (por ejemplo, muchos libros)? Como observará el lector, a la hora de construir el instrumento no se pretendía responder a esta pregunta, sino que el objetivo simplemente era «conocer» los hábitos lectores, en tanto que describirlos y cuantificarlos en la población de adolescentes. Si el objetivo hubiera sido «medirlos», sería necesario establecer una regla aún más precisa, según la cual, las distintas conductas o indicadores observables estén expresados en una misma dirección, de tal forma que, como decíamos anteriormente, se puedan «sumar» sus efectos para determinar la «cantidad» diferencial que corresponde a cada sujeto en función de su conducta.

En esta línea, consideramos que los instrumentos que permiten la medición son aquellos que, además de estar fundamentados en un sistema teórico que se concreta en un conjunto de indicadores observables a evaluar, se basan en una regla desde la cual resulta posible sumar los efectos por separado de cada indicador, a fin de conocer la cuantía del atributo medido que posee un objeto o sujeto[12]. Esto hace que el instrumento sea capaz de cuantificar las diferencias entre los valores de atributo que poseen los distintos sujetos. La construcción de esta regla, por otra parte, requiere de elaboraciones y operativizaciones teóricas muy exhaustivas, rigurosas y detalladas, de tal forma que exista un mínimo consenso entre la comunidad científica de que la regla construida es válida y podemos usar con garantías la medida que nos ofrece.

Bajo este criterio, consideramos que sólo los instrumentos que vamos a ver en este tema (tests, escalas de actitudes y pruebas objetivas) permiten la medición, además de

algunos sistemas de categorías que están construidos de forma que las conductas que observan se pueden considerar indicadores de efectos aditivos en la evaluación de un determinado rasgo o característica del comportamiento. En nuestra opinión, instrumentos como las listas de control, las escalas de observación, determinados sistemas de categorías, los cuestionarios y las entrevistas estructuradas *no* permiten la medición, en los términos que hemos venido defendiendo.

No obstante, es necesario matizar esta afirmación. A pesar de nuestro esfuerzo por diferenciar conceptos, entendemos que los límites entre la cuantificación y la medición en sentido estricto no están claramente definidos. Así, determinados cuestionarios (o listas de control, etc.) pueden llegar a aproximarse bastante a la medición. El ejemplo más claro que se nos ocurre al respecto es la diferencia existente entre un test de aptitudes y un test de intereses. El primero de ellos se considera como un test en el sentido más ortodoxo del término; en él se presentan tareas a los sujetos que pueden ser resueltas «bien» o «mal» porque tienen una respuesta correcta o incorrecta. Por ejemplo, si pedimos a alguien que sume 2 y 2, sólo cuando obtenga como resultado 4 podremos decir que ha realizado la tarea convenientemente. En este caso, es fácil identificar tareas similares y elaborar una regla de medida: cada tarea bien puede valer un punto y así se pueden sumar sus efectos.

Los tests de intereses, por ejemplo los de intereses profesionales, pueden no ser considerados tests en sentido estricto. En ellos no hay respuestas «correctas» o «incorrectas», ¿o es que acaso preferir trabajar con personas es mejor o peor que preferir trabajar con objetos? En este caso, elaborar e identificar la regla mediante la cual se va a poder afirmar que una persona tiene más interés que otra en tal o cual actividad es una tarea más compleja. Sin embargo, en nuestra opinión, también en estos casos, resulta posible la medición, aunque la elaboración del instrumento es quizá más compleja y, por lo tanto, resulta más difícil alcanzar el consenso.

Esto sucede con todos los tests que, en lugar de las aptitudes, habilidades o conocimientos de las personas, pretenden evaluar rasgos o características de su personalidad, de su forma de ser y de orientarse en la vida (tests de intereses, de valores, de rasgos de personalidad, escalas de actitudes, etc.). Por eso, a estos instrumentos[13] se les llama en ocasiones «cuestionarios» y se evita usar con ellos el término «test». En nuestra opinión, si bien en ellos, como ya hemos apuntado, resulta más difícil elaborar la regla que permite la medición y, por tanto, es más difícil llegar a un consenso sobre la validez de la misma, estos instrumentos pueden también considerarse como tests. De hecho, en el campo educativo, tienen un considerable interés porque nos permiten medir dimensiones tan importantes como los valores, las actitudes, los rasgos de personalidad, etc.

Teniendo en cuenta estas consideraciones en torno a la medida, en este capítulo vamos a centrarnos en los instrumentos que permiten la medición (tests, escalas de actitudes y pruebas objetivas). Así, se hará referencia a los conceptos de *test* y *prueba*, y a las

características que los definen. Además se presentarán algunas ideas básicas relativas a su construcción, aplicación e interpretación, dedicándose un apartado final a la revisión de las características técnicas (validez y fiabilidad) que esos instrumentos deben mostrar. Los contenidos del capítulo suponen una aproximación inicial a los instrumentos de medida en educación, no obstante, es necesario advertir al lector que, dada la amplitud de este campo, se han dejado fuera muchos aspectos en los que sería deseable profundizar. La construcción y aplicación de escalas de actitudes y los diferentes tipos de escalas, el análisis de algunas pruebas muy usadas en el campo educativo, etc. serían algunos ejemplos significativos de ellos.

1.1. La medición, la evaluación y el diagnóstico

Una vez clarificado el concepto de medida, conviene ahora que centremos nuestra atención en las relaciones que se establecen entre la medida, por un lado, y el diagnóstico y la evaluación, por otro. La medida, en contraposición a la evaluación, es un proceso en el que no se emite ningún juicio de valor sobre el objeto medido (WOLF, 1990), sino que, como hemos visto, se define como la asignación de números a las cualidades de los objetos según ciertas reglas. Esto no implica que la medida no pueda formar parte del proceso de diagnóstico y evaluación. De hecho, la evaluación y el diagnóstico son términos mucho más amplios que el de medición, de tal forma que se puede decir que el diagnóstico y la evaluación incluyen a la medición.

Por otra parte, en toda actividad de diagnóstico y evaluación, recogemos información sobre la realidad objeto de estudio. Ahora bien, no toda la información que recogemos puede entenderse como una medida en sentido estricto. Es interesante, en este sentido, la aportación de LÁZARO (1986:81), quien afirma que «el diagnóstico pedagógico debe superar el concepto de medida para, basándose tanto en ella como en valoraciones cualitativas, realizar una evaluación de todo el entramado escolar, bien se refiera a producto, procesos, alumnos o institución».

En la revisión de enfoques de diagnóstico que se presenta en el capítulo 1 de este libro, se puede observar que algunos de los modelos tradicionales han identificado el diagnóstico (y la evaluación) con la medición y la testación. Es decir, partían de la consideración de que todo proceso de diagnóstico consiste en la aplicación de tests, ya que sólo estos instrumentos permiten la medición objetiva del comportamiento humano. Esto es una visión reduccionista de la actividad diagnóstica, en tanto que ésta se limita a la aplicación de tests, con las ventajas y limitaciones que estos instrumentos puedan tener.

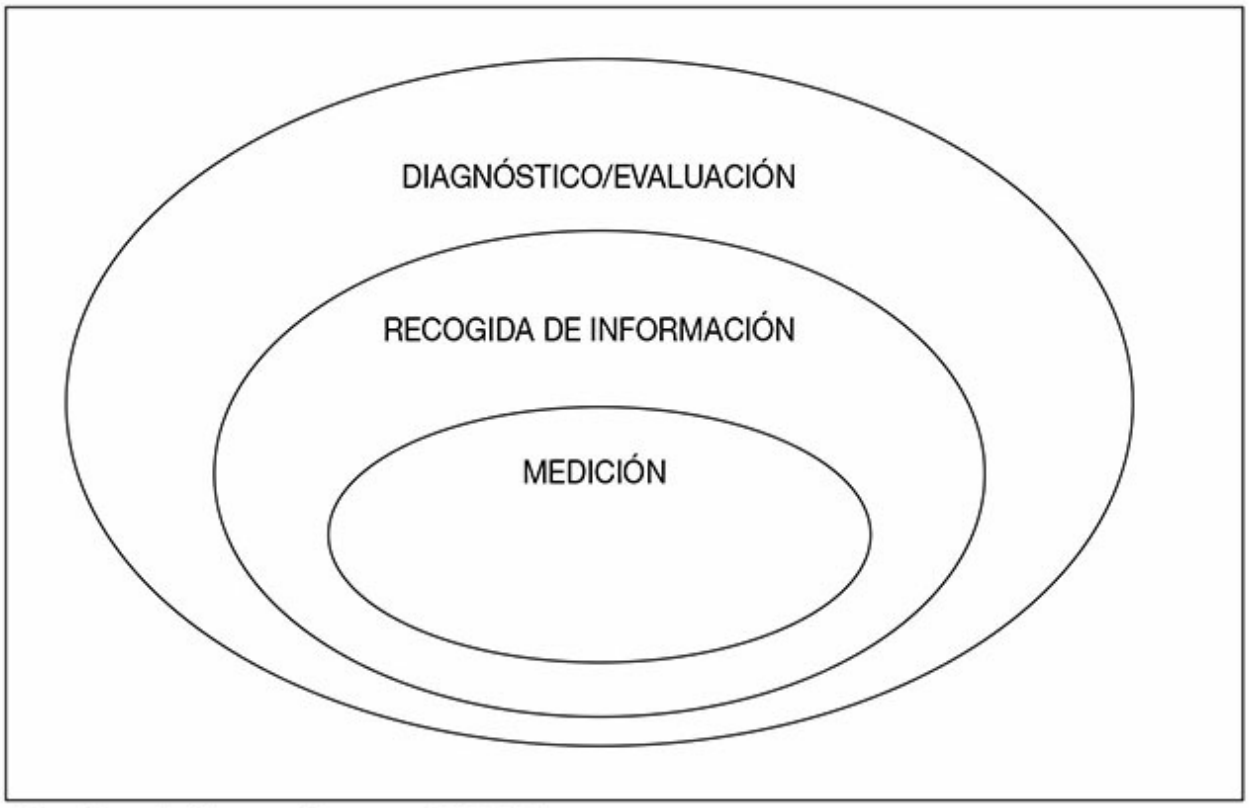
Es evidente que la medida posibilita la precisión de los conceptos, facilita la clasificación más allá de categorías gruesas y, sobre todo, es un instrumento inestimable en el momento clave de la validación de hipótesis (PÉREZ JUSTE, 1991). Sin embargo,

otro tipo de descripciones (cuantitativas y cualitativas) son necesarias para la formulación de juicios de valor y la toma de decisiones que todo diagnóstico conlleva.

En nuestra opinión, el proceso de diagnóstico y evaluación es más amplio, comprensivo e inclusivo que el de medición, en tanto que engloba otras actividades además de ésta. Así, un diagnóstico puede basarse o no en la medición, y puede partir o no de la aplicación de tests. En la figura 1 presentamos un esquema que sintetiza la relación existente entre diagnóstico/evaluación, recogida de información y medida.

FIGURA 1

RELACIÓN ENTRE DIAGNÓSTICO/EVALUACIÓN, RECOGIDA DE INFORMACIÓN Y MEDIDA



(Basado en De Ketele y Roegiers, 1995: 97)

2. LOS TESTS: CONCEPTO Y CARACTERÍSTICAS

Una primera aproximación al concepto de test proviene del hecho de que se trata de un instrumento que intenta medir el comportamiento humano. Esto implica y exige, como se ha visto en apartados anteriores, un proceso cuidadoso de elaboración. «Nada más lejos de un test que un instrumento rápidamente concebido, inmediatamente materializado y prontamente puesto al alcance del público» (PÉREZ JUSTE, 1991:124). Conseguir operativizar las características a medir en indicadores observables y traducir éstos a tareas concretas, más o menos similares, que nos permitan aproximarnos lo más posible a una unidad de medida, no es un objetivo que se consiga fácilmente ni por casualidad, sino sólo como el fruto de muchas horas de trabajo y de muchos estudios previos.

En esencia, un test es una prueba. Constituye una especie de estímulo o reactivo ante el cual se pone a prueba a los sujetos para que éstos evidencien o manifiesten algunas de sus características o conductas (la posesión de una habilidad, el dominio de un conocimiento...). En esta definición de test, algunos autores (KERLINGER, 1975; PÉREZ JUSTE, 1989), resaltan que, al ser una prueba y representar un cierto desafío, quedan excluidos determinados instrumentos de medida como son las escalas y pruebas para la evaluación de la personalidad, intereses, valores, etc.

Para ANASTASI (1982:21), un test «constituye esencialmente una medida objetiva y tipificada de una muestra de conducta». Según esta autora, las pruebas o tests psicológicos son como las pruebas en cualquier otro campo científico. En ellas, las observaciones se realizan sobre una muestra pequeña, pero cuidadosamente elegida de la conducta de un sujeto.

La tipificación en los tests se refiere a la uniformidad de procedimiento en la aplicación y puntuación del test (ANASTASI, 1982). Así, si tenemos intención de comparar las puntuaciones obtenidas en una determinada prueba por varios sujetos, es necesario garantizar que las condiciones de aplicación del test van a ser las mismas para todos ellos. Un factor que va a propiciar la tipificación es la formulación detallada de las instrucciones de aplicación de la prueba. Asimismo, otro paso necesario para la tipificación será el establecimiento de normas que permitan la interpretación de las puntuaciones que arroja una prueba. Estas normas se refieren al nivel de ejecución medio o *normal* de otros sujetos.

Algunos autores, como se observa en la definición de ANASTASI (1982), señalan el carácter objetivo de la medida que proporcionan los tests. La forma en que se aplican, la forma en que se puntúan y la forma en que se interpretan (de acuerdo con normas), factores éstos que propician su tipificación, también dotan de objetividad a los tests, en tanto que son independientes (aplicación, puntuación e interpretación) del juicio subjetivo del examinador. Una misma persona, independientemente de quien sea su

examinador, debería obtener una misma puntuación en un test. Según ANASTASI (1982), otros factores contribuyen a la objetividad de los tests; estos factores son su fiabilidad y validez, que pueden conocerse a través de procedimientos empíricos y objetivos.

MEHRENS y LEHMANN (1982) proporcionan una definición bastante comprensiva de lo que son los tests o pruebas estandarizadas, en la que, a diferencia de otras definiciones, resulta posible incluir tanto los tests referidos a normas como los referidos a criterios, y tanto los tests de aptitudes y rendimientos como los cuestionarios o inventarios estandarizados que se usan para la evaluación de la personalidad. Los autores citados basan su definición en el concepto de estandarización, como característica más sobresaliente de los tests e instrumentos de medida. Según estos autores, las pruebas estandarizadas ofrecen métodos para conseguir muestras de conducta mediante *procedimientos uniformes*. Por procedimiento uniforme, se entiende la administración de una misma serie fija de preguntas con las mismas instrucciones y limitaciones de tiempo, y un procedimiento de calificación cuidadosamente definido y uniforme. Únicamente las pruebas objetivas que no llegan a estandarizadas —en los términos descritos— quedarían fuera de la definición de MEHRENS y LEHMANN (1982).

A fin de caracterizar mejor los tests y las pruebas objetivas, PÉREZ JUSTE (1989, 1991) destaca un conjunto de notas o características propias de estos instrumentos:

a) Su elaboración y construcción. PÉREZ JUSTE (1991) resume en pocas palabras lo que significa construir un test: implica un conocimiento profundo del rasgo a medir, de las diferentes teorías en torno al mismo; la capacidad de elaborar una definición que después habrá que operativizar sobre tal rasgo o dimensión; la materialización de esta definición en una amplia gama de elementos, preguntas o cuestiones; su aplicación experimental a grupos de sujetos; la prueba de los resultados y el análisis de los mismos con vistas a su mejora; la determinación de ciertas características técnicas para evidenciar, fundamentalmente, que mide ese rasgo concreto y no otro, y que lo mide con notable precisión.

b) Su aplicación. Puesto que los tests son medidas tipificadas del comportamiento humano, es necesario asegurar la igualdad de condiciones de los distintos sujetos a la hora de realizar la prueba. En caso contrario, existirían dificultades para su correcta interpretación, ya que no podríamos estar seguros de que una diferencia en puntuaciones no es debida a factores extraños (condiciones adversas en la realización de la prueba, mayor o menor nivel de ayuda por parte del examinador...).

c) Su valoración. Los tests se caracterizan por ser unos instrumentos que pretenden alcanzar el máximo grado de objetividad posible en la medición. Esto se ve favorecido por el hecho de que se valen de respuestas concretas y de que ofrecen instrucciones tan precisas a los examinadores que se evita toda posible discrepancia entre ellos.

d) Sus características técnicas. A cualquier instrumento de evaluación se le exige que sea preciso y válido para los fines que se marca. No obstante, sólo los tests suelen incorporar sistemáticamente referencias expresas a estas características, así como a los procedimientos aplicados para su estimación. La publicación de un test se realiza sólo cuando existe este tipo de estudios empíricos que respaldan la calidad de la medida ofrecida por el instrumento, requisito éste que no siempre se da en otros tipos de instrumentos.

e) La interpretación de sus puntuaciones. Aunque, como veremos con posterioridad, existen tests y pruebas en las que la interpretación se realiza en función de criterios, una gran mayoría de los tests utilizan la interpretación basada en normas o baremos que permiten relativizar las puntuaciones de un sujeto y compararlas con las de un grupo de referencia, en lugar de quedarse en la mera descripción de los datos. Así, la valoración que se hace del nivel de ejecución de una persona no responde al criterio subjetivo del examinador (que es lo que sucede cuando decimos: «más de cinco preguntas correctas equivale a un aprobado»), sino a la comparación que se establece entre la ejecución del sujeto y la de otros sujetos de características similares que se consideran representativos de la población.

Esta caracterización general de los tests nos permite, por el momento, aproximarnos a su conceptualización. Sin embargo, algunas de estas notas definitorias propuestas por PÉREZ JUSTE (1989, 1991) no se adaptan plenamente a las características de ciertos tipos de tests y pruebas existentes en el mercado. Pero dejaremos esta cuestión para el apartado referido a los tipos de tests. Aunque en el citado apartado haremos referencia a distintas clasificaciones de estos instrumentos, de acuerdo con varios criterios, consideramos necesario plantear aquí algunas cuestiones sobre ciertos tipos de tests que entran más en el ámbito de su conceptualización que en el de su clasificación.

Así, dentro de los instrumentos dirigidos a evaluar el rendimiento (escolar o profesional), consideramos necesario diferenciar entre *tests* y *pruebas objetivas*. En general, hablamos de tests cuando la prueba en cuestión está estandarizada y tipificada. Así, los tests se construyen, generalmente, por los editores de pruebas y se destinan al uso en una gran variedad de situaciones (por ejemplo, todas las escuelas de un país, si se trata del rendimiento escolar, o todos los puestos de trabajo relacionados con la administración, si se trata de un test de rendimiento para personal administrativo). Por lo general, los tests de rendimiento suelen tener un contenido bastante amplio, ya que se enfocan a la aplicación en una amplia variedad de situaciones. Por la misma razón, suelen además presentar normas y baremos de gran alcance (ámbito nacional, por ejemplo).

Por su parte, las pruebas objetivas se construyen con un planteamiento menos ambicioso. El ejemplo más habitual es el de los docentes, que las utilizan como medio para la evaluación del aprendizaje de sus alumnos, aunque en el terreno escolar también podríamos encontrar ejemplos en los que las pruebas vienen elaboradas por un equipo

docente o, incluso, por varias escuelas de una comunidad. En el terreno laboral, podemos también encontrar ejemplos semejantes. La cuestión está en que el área de contenido será mucho más reducida; limitada al temario impartido por el profesor, al proyecto curricular de una materia, a las funciones propias de un determinado puesto laboral en una determinada empresa... Asimismo, puesto que la prueba cubre un universo mucho más estrecho, las materias se cubrirán de forma más detallada (BROWN, 1980). Las pruebas no suelen presentar baremos o normas interpretativas puesto que la finalidad con la que se construyen, la mayoría de las veces, es para aplicaciones de más restringido alcance.

No obstante, y según PÉREZ JUSTE (1989), la prueba objetiva, como *prueba* que es, participa de las características apuntadas para los tests: es un tipo de reactivo que trata de poner de relieve el dominio por parte del sujeto de determinados objetivos educativos. Como *objetiva*, indica que a través de ella se pretende llegar a puntuaciones que en nada dependen de juicios subjetivos. Por otra parte, incluyen una gran cantidad de ítems o elementos dada la rapidez y brevedad de las respuestas solicitadas a los sujetos, lo que permite un muestreo suficientemente amplio y representativo del área a evaluar.

Sintetizando, los elementos que diferencian a tests estandarizados y pruebas objetivas se pueden resumir en los siguientes (PÉREZ JUSTE, 1989):

— Aunque en ambos casos el proceso de elaboración es cuidadoso, en los tests el detenimiento y costo son mayores.

— El contenido suele ser más amplio en los tests, aunque esto no implica un mayor número de ítems.

— Por lo general, en las pruebas no existen una reglas tan rigurosas y específicas de aplicación estándar.

— Habitualmente, las pruebas no suelen acompañarse de normas de interpretación de los resultados.

Debemos aclarar también las diferencias existentes entre los tests de aptitudes e incluso de rendimiento y lo que se denominan como *cuestionarios o inventarios estandarizados* para la evaluación de dimensiones, rasgos o factores relacionados con la personalidad. En la introducción de este capítulo, ya indicamos que no todos los atributos que tienen interés desde un punto de vista educativo pueden acogerse a un formato de pregunta cuya respuesta puede ser clasificada de «correcta» o «incorrecta». Por ejemplo, entre tener tendencia a la introversión o a la extroversión, ¿cuál de los dos extremos es el correcto?, o en el campo de los intereses vocacionales, ¿qué es mejor: preferir las actividades al aire libre o en ambientes de interior? Como el lector puede pensar, este formato de pregunta se aleja mucho de demostrar que se tiene o no un conocimiento o una habilidad. Por lo tanto, propiamente hablando, los instrumentos que se proponen medir aspectos como los valores, las actitudes, los intereses y la personalidad en general no pueden llamarse «tests», en tanto que esta palabra implica

«poner a prueba» al sujeto para inferir la existencia de un rasgo. Es por ello, que habitualmente se presentan como «cuestionarios (o inventarios) estandarizados» para la evaluación de un determinado rasgo o característica.

A pesar de que somos conscientes de la importante diferencia que este tipo de instrumentos presenta, nosotros los vamos a incluir aquí dentro de la denominación genérica de tests. En cierta medida, consideramos que si bien en ellos no se pone al sujeto a prueba en el sentido de que su respuesta pueda o no ser acertada, sí se le pone a prueba indirectamente en tanto se le ofrece un reactivo ante el cual el sujeto emitirá una respuesta que podrá clasificarse dentro de un patrón determinado con vistas a su puntaje e interpretación. Con un ejemplo apreciaremos mejor esta idea.

En el SIV o *Cuestionario de Valores Interpersonale*s de GORDON (1995), se pretende evaluar la intensidad relativa de seis valores interpersonales (Estímulo, Conformidad, Reconocimiento, Independencia, Benevolencia y Liderazgo). En los distintos reactivos o ítems, se presentan al sujeto frases en tríadas que representan los diferentes valores interpersonales medidos por la prueba. Un ejemplo es:

a) Poder gobernar mi propia vida.

b) Gozar del afecto de los demás.

c) Ser comprensivo con los que tienen algún problema.

Ante cada una de las 90 afirmaciones —presentadas en 30 tríadas—, el sujeto debe señalar cuál considera como la más y la menos importante de las tres afirmaciones. Es evidente que conforme vaya respondiendo, va a ir eligiendo un conjunto de frases como representativas de unos valores o de otros, por lo que su puntuación final en cada valor se compone de la respuesta que ha ido dando a cada ítem del test. De tal forma que aquellos sujetos que hayan elegido como más importantes frases parecidas a la a) del ejemplo, alcanzarán una puntuación más alta que los demás en el valor Independencia. Como puede apreciarse, si bien la respuesta de los sujetos no puede calificarse como «correcta» o «incorrecta», sí es posible calificarla dentro de un patrón (Independencia, Estímulo, Conformidad, etc.).

En todo lo demás, el proceso de elaboración, aplicación e interpretación de este tipo de pruebas se asemeja al de cualquier test de aptitudes o de rendimiento. Así, al igual que en un test de aptitudes o rendimiento resulta complejo plantear una tarea que sea verdaderamente indicativa de la aptitud o dominio que se pretende medir, en un cuestionario de este tipo también habrá que poner especial cuidado en la redacción de los ítems, ya que éstos deben ser, al igual que en el otro caso, buenos indicadores del valor, interés o rasgo en general, que se pretende evaluar.

Por esta razón, en este texto los vamos a considerar dentro del grupo genérico de tests y pruebas objetivas. Por un lado, por las razones expuestas, su proceso de elaboración, aplicación e interpretación requiere de las mismas tareas y precauciones que el resto de pruebas de este grupo. Por otro, constituyen instrumentos muy valiosos para la

intervención educativa, pues permiten evaluar aspectos y dimensiones que, en dicha intervención, no se deben soslayar.

Hecha esta breve conceptualización de los tests como instrumentos de medida, conviene que nos detengamos a pensar en las funciones para las que se utilizan. Como consecuencia de sus características (objetividad, tipificación, etc.), los tests tienen múltiples aplicaciones tanto en el campo pedagógico como en otras disciplinas sociales y humanas. Al igual que otros instrumentos, son herramientas que permiten la realización de diagnósticos tanto individuales como colectivos y, a partir de cuyos resultados se pueden extraer implicaciones importantes para la toma de decisiones educativas y la intervención orientadora. Asimismo, son unos instrumentos muy valiosos para la medición de variables en la comprobación de hipótesis científicas.

Sin embargo, además de estas aplicaciones, los tests se han usado para otros fines más controvertidos: clasificar sujetos en categorías a fin de administrarles un determinado programa o adscribirlos a una determinada modalidad educativa; seleccionar a los sujetos para su admisión en distintas instituciones (para un trabajo, para entrar en una Universidad...), etc. Estas aplicaciones se basan en su supuesta capacidad para predecir el comportamiento futuro de los sujetos y es lo que algunos autores (McDIVITT y St. JOHN, 1996:57) han denominado como «el test-portero» o «el guardián de la puerta», en tanto que las puntuaciones obtenidas por un sujeto en una prueba «le abren o no la puerta» para acceder a una determinada institución, programa educativo o puesto laboral. A nuestro juicio, los tests tienen muchas limitaciones como para que este tipo de predicciones o pronósticos puedan considerarse válidos. De esta forma, entendemos que este tipo de comportamientos responde a una confianza excesiva en la capacidad de los tests, lo cual entra en lo que en el capítulo 1 entendíamos como un mal uso o mala práctica de estos instrumentos.

Es evidente que cuestiones como el rendimiento escolar o el rendimiento profesional dependen de muchas variables, por lo que ¿cómo pueden predecirse con éxito sobre la base de las puntuaciones de un test, por bien construido que esté el instrumento? Eso sin olvidar que el «rendimiento futuro» no es la única cuestión que está en juego; también son importantes factores como la propia satisfacción personal del sujeto con lo que hace. Asimismo, es posible que los resultados de un sujeto en una determinada prueba indiquen que ese sujeto va a beneficiarse escasamente de un determinado programa educativo, pero, desde una perspectiva pedagógica, ¿puede considerarse adecuado que como no se beneficiará «todo lo deseable» es mejor que «no se beneficie en nada»?

En lugar de una predicción tajante, los resultados de los tests deben ofrecer al sujeto una descripción realista de las dificultades y ventajas que puede encontrar a la hora de seguir un determinado curso de acción, siempre sobre la base de la falibilidad del instrumento desde el que realizamos esta descripción. Como plantea ANASTASI (1992:612), «las puntuaciones de un test nos dicen cuál es el rendimiento de las personas en el momento de la prueba, pero no por qué rinden así». Además, ninguna decisión

educativa importante debe tomarse sobre la base de una única prueba (MEHRENS y LEHMANN, 1982). Precisamente, entre los estándares de prestigiosas asociaciones americanas, se encuentra el siguiente: «Tanto en la educación primaria como secundaria, cualquier decisión o caracterización que pueda tener un impacto importante en la persona que cumplimenta el test no debe ser automáticamente tomada sobre la base de la mera puntuación en el mismo. Los profesionales que tienen que decidir deben tomar en consideración otra información relevante para la decisión» (AERA/APA/NCME, 1985:54).

Como ha sugerido MESSICK (1980), los tests no sólo deben evaluarse en términos de sus propiedades psicométricas, sino que también las aplicaciones de los tests deben ser evaluadas en función de sus potenciales consecuencias sociales. De esta forma, justificar el uso de tests exclusivamente en las evidencias empíricas de su validez no es suficiente; en opinión de este autor, es necesario también analizar las potenciales consecuencias sociales de su uso y no sólo desde un planteamiento de costes y beneficios, sino también en términos de sus posibles efectos colaterales.

3. TIPOS DE TESTS

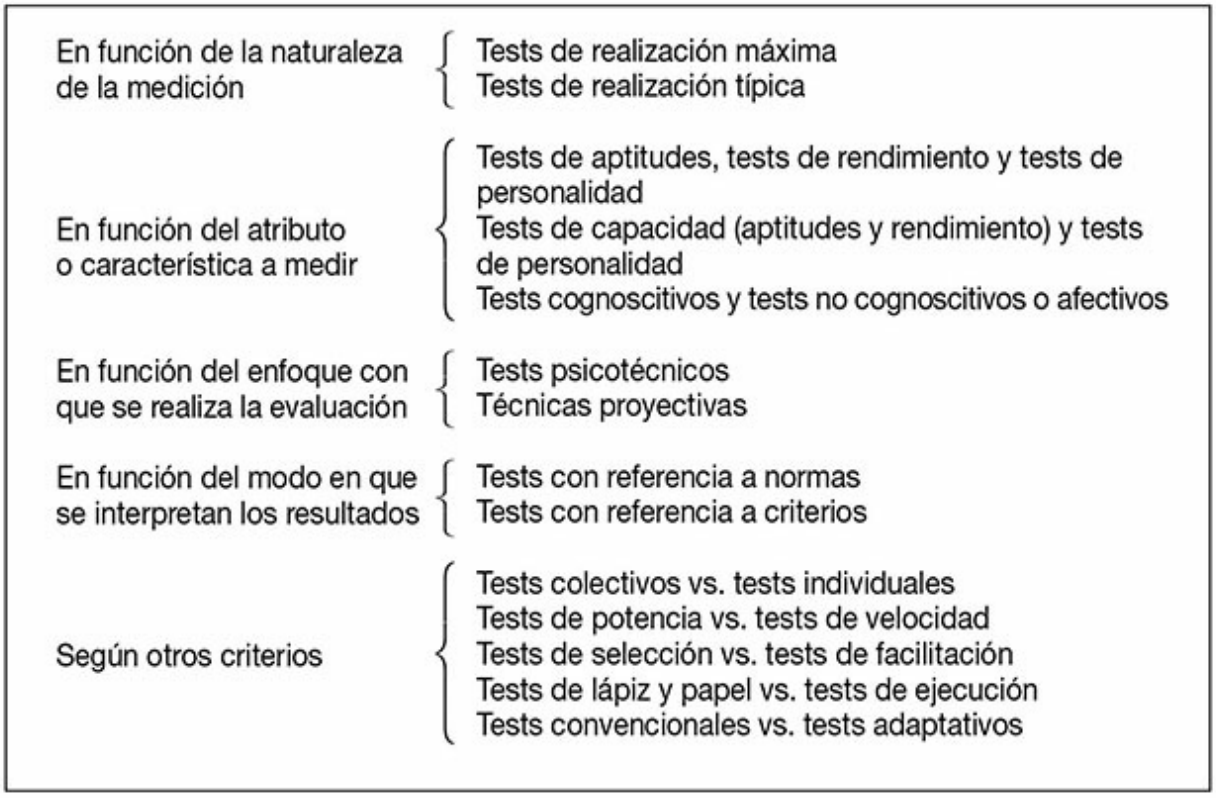
Existen muchos tipos diferentes de tests. En el apartado anterior, de hecho, ya hemos mencionado algunas diferencias importantes en estos instrumentos de medida. A lo largo de este apartado, no obstante, presentaremos varias clasificaciones de tests, realizadas con diversos criterios, lo que nos permitirá, además de conocer las distintas modalidades que existen, adentrarnos con más profundidad en este tipo de instrumentos, en función de las diferentes formas y características que pueden adoptar.

En primer lugar, presentaremos algunas clasificaciones de tests en función de aspectos tan importantes como son la naturaleza de la medición, los contenidos o características que evalúan, la forma de enfocar la evaluación y la forma de interpretar los resultados. Asimismo, se mencionan algunas clasificaciones de menor relevancia pero que contribuyen a conocer la diversidad de formas que los tests pueden tomar. Todas estas clasificaciones y las categorías que engloban aparecen sintetizadas en la figura 2.

Posteriormente, presentamos un intento de clasificación temática de los tests más usados en el campo de la orientación. Esta clasificación, que puede considerase un mini-vademécum, contiene una relación de gran parte de las pruebas comercializadas. A partir de ella, el lector interesado puede, además de conocer las alternativas existentes, acceder al instrumento que le interese.

FIGURA 2

CRITERIOS DE CLASIFICACIÓN DE LOS TESTS Y TIPOS DE TESTS


En función de la naturaleza de la medición
Tests de realización máxima
Tests de realización típica
En función del atributo o característica a medir
Tests de aptitudes, tests de rendimiento y tests de personalidad
Tests de capacidad (aptitudes y rendimiento) y tests de personalidad
Tests cognoscitivos y tests no cognoscitivos o afectivos
En función del enfoque con que se realiza la evaluación
Tests psicotécnicos
Técnicas proyectivas
En función del modo en que se interpretan los resultados
Tests con referencia a normas
Tests con referencia a criterios
Según otros criterios
Tests colectivos vs. tests individuales
Tests de potencia vs. tests de velocidad
Tests de selección vs. tests de facilitación
Tests de lápiz y papel vs. tests de ejecución
Tests convencionales vs. tests adaptativos

3.1. Algunas clasificaciones de tests aplicables al campo educativo

Los criterios mediante los cuales se pueden clasificar los tests son muchos. Podemos hablar de tests verbales y no verbales, de tests de tarea única y baterías de tests, etc. En este apartado vamos a revisar algunos de los criterios que consideramos más relevantes a la hora de analizar, diferenciar y clasificar los tests. La mayoría de estos criterios no son excluyentes entre sí, por lo que nada se opone a que combinemos varios de ellos para caracterizar una prueba concreta. Así, por ejemplo, una prueba puede ser un test de tendencia psicotécnica, de realización máxima, de aptitudes y con referencia a normas. Además de colectiva, de lápiz y papel, de velocidad, etc.

Empezando ya con la presentación de cada uno de estos criterios, una de las principales características que podemos tener como referente a la hora de clasificar los tests se refiere a la **naturaleza de la medición** que proporcionan. Bajo este criterio, CRONBACH (1972) establece una diferenciación entre dos grandes categorías que ha sido ampliamente utilizada: por un lado, los tests de ejecución máxima y, por otro, los de ejecución típica.

Los *tests de ejecución máxima* buscan medir la capacidad máxima que un sujeto puede demostrar en su ejecución. En ellos, los sujetos tratan de obtener la máxima calificación, ya que la meta es medir los límites de su capacidad (BROWN, 1980). Las instrucciones de la prueba suelen indicar al sujeto que ha de intentar obtener la puntuación más alta posible. En general, las pruebas de aptitudes y de rendimiento son tests de este tipo.

A diferencia de los anteriores, los *tests de ejecución típica* se enfocan a evaluar la conducta habitual y típica de los sujetos en determinadas situaciones, por lo que en sus instrucciones se insiste, sobre todo, en que el sujeto exprese sus reacciones habituales. En este caso, no se trata de poner a prueba lo que una persona puede y sabe hacer, sino lo que ésta suele hacer, es decir, sus reacciones y conductas habituales. Son medidas de ejecución típica las pruebas de personalidad, intereses y actitudes.

El **atributo o característica medida** a través del test es otro criterio muy utilizado para clasificar estos instrumentos. Bajo este criterio se suele diferenciar entre tests de personalidad, de aptitudes y de rendimiento. Dentro de los *tests de personalidad* vendrían a estar las mismas pruebas ya mencionadas para los tests de realización típica, es decir, aquellas que, abarcando una cierta diversidad de contenidos, miden factores que habitualmente se agrupan bajo el término *personalidad*: los valores, las actitudes, los intereses, la adaptación persona-medio, los rasgos del carácter, etc. Dado que en este tipo de pruebas, como ya se ha indicado, las respuestas no pueden ser consideradas como verdaderas o falsas, se suelen preferir los términos *inventario* o *cuestionario* para referirse a ellas.

Respecto a los otros dos grupos mencionados (tests de aptitudes y rendimiento), en ocasiones aparecen clasificados en un grupo más amplio denominado «pruebas de capacidad» (CABRERA y ESPÍN, 1986 y THORNDIKE y HAGEN, 1978). Así, las pruebas de capacidad (que pueden ser de aptitudes o de rendimiento) pretenden medir lo que puede o podría hacer el sujeto en función de sus potencialidades personales y del rendimiento adquirido en un aprendizaje determinado. A veces, a las pruebas de capacidad se las denomina genéricamente como *tests cognoscitivos*, a diferencia de las pruebas de personalidad que reciben el nombre de *medidas no cognoscitivas o afectivas* (MEHRENS y LEHMANN, 1982).

Los *tests de aptitudes* se enfocan a medir capacidades para adquirir ciertas conductas o habilidades, dadas las oportunidades adecuadas. Es decir, tienen como fin medir lo que se puede aprender, mediante un entrenamiento apropiado (BROWN, 1980). La habilidad a aprender puede variar desde una habilidad intelectual compleja (como, por ejemplo, deducir relaciones abstractas) hasta actos motrices simples. Por su parte, las *pruebas de rendimiento* miden los resultados del aprendizaje (CABRERA y ESPÍN, 1986).

Esta diferenciación que puede parecer clara a primera vista, plantea algunos interrogantes, ya que en muchas ocasiones resulta difícil diferenciar un test de aptitud de un test de rendimiento. En general, y como plantea GARCÍA JIMÉNEZ (1997), las pruebas

de rendimiento vendrían a evaluar la eficacia de actuaciones específicas, académicas o profesionales, para la consecución de determinados objetivos de aprendizaje (aprender a leer, a calcular, conocer la geografía, aprender mecanografía, etc.). Sin embargo, los objetivos educativos no se refieren al dominio de conceptos en exclusiva, sino que abarcan la adquisición de procedimientos, habilidades y destrezas. Es por ello que resulta difícil diferenciar con nitidez las pruebas dirigidas a evaluar rendimientos y las dirigidas a evaluar aptitudes. Al respecto, BROWN (1980) propone dos criterios que nos permiten diferenciar los tests de rendimiento de los de aptitudes. Así, un test es de rendimiento si mide el aprendizaje que se ha producido: 1) como consecuencia de las experiencias en una situación de aprendizaje relativamente circunscrita (un programa de estudios, un curso de entrenamiento,...) y 2) cuando el marco de referencia está en el presente o el pasado, o sea, en lo que se ha aprendido.

Muchos autores (BROWN, 1980; THORNDIKE y HAGEN, 1978) han defendido que una de las diferencias más notorias entre los tests de aptitudes y los de rendimiento se refiere a que los primeros se enfocan a predecir o pronosticar la ejecución futura del sujeto (función predictiva de los tests de aptitudes). Ahora bien, como plantea GARCÍA JIMÉNEZ (1997), hoy día parece claro que uno de los predictores más eficaces del rendimiento futuro, tanto académico como profesional es, sin duda, el rendimiento previo. A ello podríamos añadir que, aunque con los tests de aptitudes nos interese conocer el potencial de desarrollo futuro, con ellos al fin y al cabo, no hacemos sino estimar la *ejecución actual*. Por lo tanto, la diferencia no radica tanto en lo que hacen las pruebas, sino en el propósito de sus autores y en el método que utilizaron en su elaboración (MEHRENS y LEHMANN, 1982).

Otro posible criterio de diferenciación entre ambos tipos de tests se refiere al **contenido de las tareas** que utiliza la prueba en cuestión. Así, si estas tareas están estrechamente relacionadas con el aprendizaje formal, parecería más claro y evidente clasificar la prueba como de rendimiento. No obstante, este criterio dejaría fuera todas aquellas pruebas que miden aprendizajes profesionales y no escolares. Por lo que preferimos quedarnos con el criterio apuntado por MEHRENS y LEHMANN (1982), a pesar de lo genérico que pueda resultar. La intención del autor que la elabora parece el único criterio claro a la hora de denominarla como de rendimiento o aptitud, teniendo presente que en las distintas aplicaciones que se hagan de la misma, la finalidad puede ser otra. Quizá, la cuestión está, como ya planteó CRONBACH (1972), en situarlas en una especie de continuo en el que el rendimiento o la aptitud serían los dos extremos. Mientras mayor sea la orientación de una prueba hacia un contenido determinado y más dependiente sea de la enseñanza, más se acercará a un test de rendimiento y menos a uno de aptitud.

Asimismo, dentro de cada uno de los tres grandes grupos presentados (tests de aptitudes, rendimiento y personalidad) se pueden establecer nuevas divisiones en las que no profundizaremos por considerar que escapan al objetivo de esta obra. Baste decir que los tests de aptitudes, en función de los diversos modelos imperantes en torno a la

capacidad intelectual, pueden enfocarse a medir la inteligencia como capacidad general de la persona para manejar ideas y entender relaciones abstractas (tests de factor «g» o de inteligencia general) o pueden dirigirse a evaluar diferentes sectores más restringidos de las capacidades humanas (tests de aptitudes especiales). Por su parte, y como ya se ha dejado entrever, los tests de rendimiento pueden medir aprendizajes académicos y escolares o aprendizajes de tipo profesional. Dentro de cada grupo se podrían establecer nuevas diferencias en función del contenido de aprendizaje concreto que evalúan[14]. Por último, dentro de los tests de personalidad, algunos se enfocan a la evaluación de rasgos de la personalidad normal (afabilidad, apertura al cambio, perfeccionismo, etc.), otros a las características de la personalidad anormal o patológica (tendencias paranoicas, neuróticas y psicopáticas), otros se centran exclusivamente en los valores, o los intereses, etc. Además, muchos autores (BROWN, 1980; CABRERA y ESPÍN, 1986; THORNDIKE y HAGEN, 1978; etc.) incluyen en el grupo genérico de pruebas de personalidad las escalas de actitudes, de gran aplicabilidad en el campo educativo.

Un nuevo criterio de clasificación, el **enfoque de la evaluación,** se aplica fundamentalmente a lo que antes denominamos como ámbito no cognoscitivo o afectivo. En este ámbito y bajo este criterio, pueden diferenciarse dos enfoques en la evaluación, distinguidos inicialmente por CRONBACH (1972). Más que de dos tipos de tests diferentes, podemos hablar de dos formas de conceptualizar y entender la evaluación en este campo. Así, siguiendo a CRONBACH (1972), una parte de las pruebas que evalúan la personalidad se adscriben a un enfoque psicotécnico, en tanto persiguen obtener una estimación cuantitativa de algunos aspectos aislados de la ejecución del sujeto. Otro conjunto de pruebas sigue un enfoque holístico en tanto que pretenden conseguir una visión global de la personalidad, haciendo hincapié en las dimensiones cualitativas de la conducta. En este tipo de pruebas, la atención se centra sobre el conjunto de la personalidad, en lugar de hacerlo sobre la medida de rasgos específicos.

Los tests que siguen un *enfoque psicotécnico* pertenecen a la tipología de tests a la que está dedicada este capítulo. Es decir, aquellos que se hacen sobre la base de la estandarización y la tipificación de la conducta. Presentan un diseño estructurado, con objeto de que todos los sujetos interpreten la tarea de la misma forma; las respuestas que se solicitan tienen una forma similar y condicionada para todos los sujetos a los que se aplica la prueba; pueden ser valoradas de una manera objetiva no dependiendo de la fluencia o de la habilidad expresiva del sujeto; sirven para conocer el resultado de una ejecución, no prestando excesiva atención a las demás reacciones y manifestaciones del sujeto; etc.

Por su parte, los tests que siguen un *enfoque holístico* son aquellos que reciben la denominación genérica de métodos o técnicas proyectivas. En ellos se parte de un enfoque global y no centrado en una serie de rasgos, como hacemos cuando utilizamos determinados inventarios o cuestionarios de personalidad. Las tareas que se presentan a los sujetos tienen un nivel de estructuración mínimo, de modo que se produce una mayor variación cuando el sujeto interpreta la tarea y ofrece su respuesta. Según ANASTASI

(1982), las tareas no estructuradas de las técnicas proyectivas permiten una variedad casi ilimitada de respuestas posibles (ANASTASI, 1982), aunque esto implica que sea más difícil comparar a un grupo de sujetos en un aspecto común.

En las técnicas proyectivas —de las cuales, el famoso Test de Manchas de Rorschach es el máximo exponente—, la suposición básica es la de que una persona responderá a una situación no estructurada o ambigua según su estructura de personalidad (BROWN, 1980). Por lo tanto, los estímulos (o preguntas) suelen ser casi siempre ambiguos, no estructurados o incompletos, y la tarea del sujeto es la de dar cierta estructura, un significado o una interpretación para cada estímulo. A partir de estas respuestas, el evaluador saca inferencias sobre la dinámica y la estructura de la personalidad del sujeto. Según THORNDIKE y HAGEN (1978), cuanto más vago sea el estímulo ofrecido, mayor será la oportunidad de que el sujeto se proyecte en el informe que ofrece (o proyecte en él sus motivaciones, necesidades, formas de pensamiento, etc.). Así, los tests proyectivos operan con materiales poco estructurados del tipo: un dibujo vago, una mancha de tinta, una palabra... Además las instrucciones ponen escasas restricciones al sujeto, concediéndose, en general, mayor libertad en la respuesta y no limitándose el tiempo de duración (CABRERA y ESPÍN, 1986).

Una característica importante de los métodos proyectivos es que son menos susceptibles a la falsificación y a la influencia de patrones de respuesta (como los que se verán en el último apartado de este capítulo) que suelen afectar a las medidas de la personalidad (MEHRENS y LEHMANN, 1982). Así, se asume que cuanto menos estructurados o más ambiguos sean los estímulos, menos probable es que provoquen respuestas defensivas por parte del sujeto. Aunque también es de señalar que presentan, como contrapartida, una mayor incidencia de la subjetividad del evaluador.

Los tests pueden también ser clasificados según el **modo en que se interpretan los resultados** obtenidos en ellos. Se puede distinguir, bajo este criterio, entre los tests referidos a normas y los tests referidos a criterios, menos abundantes estos últimos en el mercado editorial. La distinción entre interpretación criterial y normativa se ha utilizado sobre todo dentro del campo de los tests de rendimiento aunque, en opinión de ANASTASI (1982), en teoría sería aplicable a cualquier clase de tests. No obstante, como se ha indicado, todavía son escasas las pruebas de rendimiento que utilizan una interpretación basada en criterios.

Según BROWN (1980), en las pruebas o *Tests con Referencia a Normas* (TRN) se compara a los individuos unos con otros, por lo que la ejecución se interpreta casi siempre en función de una clasificación relativa. Por esta razón, en la interpretación basada en normas se requiere de la existencia de un grupo (*grupo normativo*) en función del cual se compara el nivel de ejecución del sujeto. Así, el sujeto ocupará una posición relativa en este grupo según su nivel de ejecución en el test, y para interpretar su puntuación diremos que «deja por debajo de sí al 85% de la población», o «está dos unidades de desviación típica por debajo de la media del grupo».

Los *Tests Referidos a Criterios* (TRC), por otra parte, ponen de relieve la destreza de contenidos y habilidades claramente definidos, o sea, lo que el sujeto sabe y puede hacer. Así, las puntuaciones de una prueba criterial se interpretan indicando qué habilidades o conocimientos ha llegado a dominar el sujeto. Esto exige una especificación muy clara y precisa de un dominio de conducta perfectamente definido (POPHAM, 1983). En una interpretación con base en un criterio se pone el acento en la descripción específica e individual que el sujeto ha realizado. En ella hacemos descripciones del tipo «compone frases negativas simples (sujeto, verbo y objeto directo) en inglés con el verbo *do* sin error», o «diferencia la secuencia de las notas musicales».

A pesar de las diferencias existentes entre ellos, ambos tipos de tests comparten un conjunto de características comunes (GRONLUND, 1974): los dos requieren especificar el dominio de rendimiento que se pretende medir; requieren una muestra relevante y representativa de ítems; utilizan los mismos tipos de ítems y las mismas reglas para su redacción y elaboración; y se valoran por las mismas cualidades de bondad (validez y fiabilidad).

No obstante, GRONLUND (1974) destaca también los elementos diferenciales de cada uno de ellos, que aparecen sintetizados en el cuadro 1. En general, puede decirse que los TRN enfatizan la discriminación entre sujetos, y los TRC subrayan la descripción de la realización de esos sujetos (GARCÍA JIMÉNEZ, 1997).

Para terminar con la presentación de clasificaciones de tests, hemos recopilado algunas de ellas que, si bien no son tan importantes y usadas como las anteriores, nos permiten establecer y matizar las diferencias entre los tests existentes en el mercado. Así, basándonos en clasificaciones aportadas por autores como ADAM (1983), CABRERA y ESPÍN (1986), GRONLUND (1974), PÉREZ JUSTE (1989), etc., podemos diferenciar los siguientes nuevos tipos de tests:

— *Tests colectivos frente a tests individuales*. La mayoría de los tests permiten una aplicación colectiva, con el ahorro de tiempo que esto implica. Sin embargo, algunos requieren la aplicación individual, bien por la complejidad de sus instrucciones, de sus tareas o bien por otras razones. El Test de Manchas de Rorschach es un buen ejemplo de este segundo grupo.

— *Tests de potencia frente a tests de velocidad.* Según PÉREZ JUSTE (1991), los tests de potencia pretenden determinar el nivel de ejecución máximo de los sujetos, por lo general, mediante una ordenación de los elementos según su dificultad, de forma que cada sujeto llega a un determinado punto del test. Los tests de velocidad, en cambio, se basan en el número de ítems que un sujeto puede completar en un tiempo dado.

CUADRO 1

ELEMENTOS DIFERENCIALES EN LOS TRN Y LOS TRC

TESTS CON REFERENCIA A NORMAS	TESTS CON REFERENCIA A CRITERIOS
Cubren un dominio extenso de tareas de aprendizaje, aunque se utiliza un número pequeño de ítems para cada tarea.	Cubren dominios de aprendizaje más específicos y delimitados a los que se dedican un número grande de ítems para cada tarea.
Ponen el acento en la discriminación entre sujetos en términos de sus diferencias relativas de nivel.	Subrayan la descripción del tipo de tareas individuales que los sujetos pueden o no realizar.
Predominan los ítems de dificultad media.	Se relaciona la dificultad de los ítems con la dificultad de la tarea, existiendo ítems de distintos niveles de dificultad.
Se utilizan fundamentalmente para examinar.	Se utilizan fundamentalmente para comprobar el dominio.
Su interpretación requiere un grupo claramente definido.	Su interpretación requiere un dominio de rendimiento claramente definido y delimitado.

(Basado en GRONLUND, 1974)

— *Tests de selección frente a tests de facilitación.* En algunos tests los sujetos se limitan a escoger la respuesta entre el conjunto de alternativas que se les presentan, mientras que en otros, tienen que elaborar y facilitar dicha respuesta.

— *Tests de lápiz y papel frente a tests de ejecución*. Los tests de lápiz y papel requieren el dominio del lenguaje oral y escrito. Sin embargo, muchas personas no lo tienen (porque no saben leer o escribir, o hablan otra lengua...). En estos casos, existen pruebas en las que no es necesario el lenguaje para la cumplimentación de una tarea (tocar objetos en el mismo orden que el examinador, disponer piezas con diferentes formas en marcos preparados, etc.).

— *Tests convencionales frente a tests adaptativos*. Los tests adaptativos o «tests a medida» son de más reciente aparición. Gracias a sofisticados análisis estadísticos, tienen la virtualidad de adaptarse a la ejecución de cada sujeto. Así, conforme éste va dando respuestas acertadas o erróneas, se le va ofreciendo una nueva tarea más ajustada a su nivel de ejecución. Generalmente, el test comienza con una tarea de dificultad media; si el sujeto la responde bien se seguirá con el mismo o mayor nivel de dificultad. En caso contrario, se buscará una tarea más fácil. Es necesario administrarlos por ordenador, dada la complejidad inherente a su aplicación y a la asignación de la puntuación. En resumen, puede decirse de los tests adaptativos que

«aun midiendo el mismo contenido no son idénticos para todos los sujetos» (GARCÍA JIMÉNEZ y GARCÍA VIDAL, 1997:148).

3.2. Clasificación temática de algunos tests y pruebas para la evaluación del desarrollo personal y profesional

Puesto que en el apartado anterior se han mostrado diferentes criterios a la hora de clasificar y agrupar los tests y los diferentes tipos de pruebas a los que estos criterios dan lugar, a continuación, vamos a presentar un conjunto más o menos amplio de pruebas, esta vez ordenadas en función del contenido (o campo conceptual) en el que cada una de ellas se sitúa. Es decir, la siguiente clasificación se enfoca a clasificar las pruebas según los constructos o rasgos que éstas aspiran a medir o evaluar.

Es importante advertir que esta clasificación no pretende ser sistemática y exhaustiva, sino que está elaborada desde un criterio específico, que consiste en recopilar precisamente aquellas pruebas que son especialmente útiles desde la perspectiva de la *orientación para el desarrollo personal y profesional.* Dado que en el campo educativo las pruebas existentes son muchas y, habiendo acotado el ámbito de esta obra en los procesos de intervención orientadora (ver los principios del diagnóstico analizados en el capítulo 1), se han dejado fuera un conjunto de pruebas muy utilizadas en el campo escolar, pruebas que por otro lado aparecen recogidas en otras clasificaciones más sistemáticas y exhaustivas (ASENSI y LÁZARO, 1979). Ello quiere decir que temáticas tan usuales en el trabajo pedagógico como la evaluación de las dificultades de aprendizaje, la evaluación del desarrollo cognitivo e intelectual, o la evaluación del nivel de competencia curricular, entre otras muchas, no han sido incluidas en la clasificación.

Teniendo presente esta aclaración, a continuación se presentan diversas categorías de pruebas, acompañadas cada una de ellas de un conjunto de tests actualmente en circulación.

3.2.1. Pruebas estandarizadas para la evaluación de los procesos de estudio y aprendizaje

Puesto que dentro del campo de la mejora del proceso de estudio y aprendizaje se han dado dos perspectivas teóricas diferenciadas, existen también instrumentos diseñados desde cada uno de estos enfoques. Así, desde la perspectiva tradicional de las técnicas de estudio, se han utilizado instrumentos estandarizados en los que se presenta al sujeto un conjunto de frases o afirmaciones respecto a las cuales debe indicar si son válidas para describir su conducta y en qué grado o intensidad. Generalmente, suelen evaluar diferentes áreas: condiciones físicas y ambientales del estudio, planificación y estructuración del tiempo y conocimiento de las técnicas básicas. Permiten una aplicación, tanto individual, como colectiva y suelen presentar baremos que permiten la comparación de la puntuación del estudiante con otros sujetos de parecidas

características profesionales. Algunos ejemplos de este grupo en el ámbito español son el CHTE o *Cuestionario de Hábitos y Técnicas de Estudio*, el IHE o *Inventario de Hábitos de Estudio* y el DIE (1, 2 y 3) *Diagnóstico Integral del Estudio*. Estas tres pruebas son distribuidas por TEA[15].

En el enfoque de las estrategias de aprendizaje también nos encontramos con este tipo de instrumentos. En ellos se presenta al sujeto un conjunto de situaciones supuestas o preguntas concretas y éste debe reflexionar y responder de acuerdo con lo que sabe sobre sus estrategias de aprendizaje. Por lo general, plantean una situación de aprendizaje y cuatro o cinco opciones sobre la forma de pensar o actuar más correcta en dicha situación. En el ámbito español disponemos de una prueba de reciente elaboración que permite la evaluación del uso de las estrategias cognitivas facilitadoras del aprendizaje: el ACRA o *Estrategias de Aprendizaje* de J. M. Román y S. Gallego (distribuida por TEA).

3.2.2. Pruebas para la evaluación de los intereses profesionales

Estos han sido durante mucho tiempo los instrumentos por excelencia en orientación profesional. Se basan en la presentación de un conjunto de profesiones y/o actividades y tareas para que el sujeto valore cuantitativamente el grado en que éstas le atraen o disgustan. Se diseñan con la finalidad de ayudar al sujeto a que elija entre distintas profesiones y actividades. De esta forma, a partir de las puntuaciones obtenidas en las diversas escalas y sub-escalas de este tipo de instrumentos, puede conocerse cuáles son los campos profesionales de mayor interés para una persona, si sus elecciones son coherentes o no, si están muy o poco definidas, etc.

En nuestro país, una gran mayoría de los instrumentos de los que disponemos en orientación profesional y que están comercializados pertenecen exclusivamente a esta categoría. Podemos citar como ejemplos el *Inventario de Intereses y Preferencias Profesionales* (IPP), el *Cuestionario de Intereses Profesionales* (CIPSA) y el *Registro de Preferencias Vocacionales* Kuder-C, todos editados por TEA. También existe una versión en castellano de la *Self-Directed Search* (SDS) de Holland y col., denominada SDS Forma R (adolescentes y adultos) y Forma E (adultos con bajo nivel formativo), comercializadas por Psymtec[16].

3.2.3. Instrumentos para la evaluación del desarrollo vocacional y profesional y la toma de decisiones

Aquí encuadramos a todos aquellos instrumentos que, diseñados desde los enfoques evolutivos de la orientación profesional, se centran en la evaluación de los constructos *madurez vocacional* y *adaptabilidad profesional* y otros relacionados con el desarrollo profesional a lo largo de la vida. Asimismo, incluimos también un conjunto de pruebas que miden aspectos relacionados con la capacidad y actitud para la toma de decisiones.

Entre las pruebas existentes en nuestro país en este interesante campo, podemos citar las adaptaciones a nuestro contexto del CDI (*Career Development Inventory*) y del CMI (*Career Maturity Inventory*), realizadas por Manuel Álvarez y un equipo de profesores de la Universidad de Barcelona, que están próximas a ser publicadas y comercializadas para su utilización. Uno de estos instrumentos, el CDI, ha sido además adaptado a nuestro contexto por ANAYA y REPETTO (1994). Además, destacamos la *Escala de Conducta Exploratoria* (ESCE) publicada por RODRÍGUEZ MORENO (1999) y el *Cuestionario de Educación para la Carrera* de PEREIRA (1992), el cual es la adaptación al castellano del QEC (*Questionnaire sur l'éducation à la carrière*). Más específicamente dirigido a personas adultas, se encuadra también aquí la adaptación al contexto español del ACCI, o *Inventario para la evaluación de las Preocupaciones en el desarrollo profesional Adulto* (IPA), publicado por PADILLA (2001).

3.2.4. Instrumentos para la evaluación de los valores personales y laborales

Este grupo reúne a un menor número de instrumentos publicados, aunque se espera que en nuestro país se produzca un avance importante en la evaluación de los valores y la importancia relativa de los roles vitales. Aunque no comercializada por el momento, FIGUERA (1996) ha elaborado la escala de *Relevancia del Rol del Trabajo* (RT). Además disponemos de dos instrumentos —editados por TEA— que nos permiten evaluar valores: *Cuestionario de Valores Interpersonales* (SIV) y el *Cuestionario de Valores Personales* (SPV). Por último, aunque no centrado específicamente en los valores, sino en las motivaciones profesionales, merece especial atención la *Escala de Motivaciones Psicosociales* (MPS), también editada por TEA.

3.2.5. Instrumentos para la evaluación de ambientes, situaciones y problemas personales y profesionales

Diseñados, por lo general, para personas adultas, se centran en cuestiones como la satisfacción profesional, el estrés, el estatus profesional, las características de los ambientes familiares, educativos y laborales, etc.

En nuestro país, podemos encontrar comercializados —concretamente, por TEA— el *Inventario «Burnout» de Masclach* (MBI), las *Escalas de Apreciación del Estrés* (EAE), la *Escala de Bienestar Psicológico* (EBP), el *Inventario de Adaptación de Conducta* IAC, el *Inventario Diferencial de Adjetivos para el Estudio del Estado de Ánimo* IDDA-EA, las *Escalas de Evaluación del Clima Social* (en el centro de trabajo, familia, centros penitenciarios y centros escolares), la *Escala de Satisfacción Familiar por Adjetivos* ESFA y la *Escala de Estrategias de Afrontamiento del Stress para Directivos* (este último instrumento está comercializado por Psymtec). Además, J. M. Peiró y su equipo (PEIRÓ et al., 1987a y 1987b) han validado el *Cuestionario General de Ambigüedad de Rol* (CGAR) y el *Cuestionario de Conflicto de Rol* (CCR), ambos aplicables al rol del trabajador en las organizaciones laborales.

3.2.6. Instrumentos para la evaluación del autoconcepto y la autoestima

Esta variable es un elemento muy importante tanto para la clarificación personal como profesional. Autores como D.E. Super han defendido que el proceso de desarrollo vocacional consiste, esencialmente, en el desarrollo y puesta en práctica del autoconcepto. Esto implica que cuando las personas eligen y desempeñan una profesión están intentando llevar a la práctica la imagen que tienen de sí mismas.

Entre las escasas pruebas de autoconcepto comercializadas en nuestro país, podemos mencionar el *Test de Evaluación del Autoconcepto Forma-5 (*AF-5) de TEA, que sirve para medir aspectos tan relevantes como el autoconcepto académico/profesional, social, físico, etc. También existe la prueba denominada SEI, o *Índice de autoestima* de Psymtec y el *Inventario de Autoestima en Medio Escolar* (IAME) comercializado por EOS[17]. El principal inconveniente que encontramos en estas pruebas es que, generalmente, nos indican el grado en que una persona tiene un alto o bajo concepto de sí misma y no cuáles son las dimensiones que componen esta percepción de sí. Es decir, tras la aplicación de estas y otras pruebas existentes, se puede concluir que el autoconcepto de una persona es positivo o negativo, pero no cómo se ve a sí misma esta persona.

3.2.7. Instrumentos para la evaluación de las habilidades sociales y el autocontrol

Incluimos en este grupo aquellas pruebas que tienen como finalidad evaluar un conjunto amplio de factores relativos a las relaciones sociales y al autocontrol personal. Como ejemplo, podemos mencionar la *Escala de Habilidades Sociales* EHS y el *Cuestionario de Autocontrol Infantil y Adolescente* CACIA, ambos editados por TEA.

3.2.8. Instrumentos complementarios en el proceso de orientación personal y profesional

En este grupo incluimos, básicamente, las pruebas de aptitudes, de rendimiento escolar y profesional y de dimensiones de la personalidad (tanto cuestionarios estandarizados como métodos proyectivos), que son, no obstante, las categorías más importantes en otras clasificaciones de tests, realizadas con criterios distintos a los que aquí estamos utilizando —como se puede observar, aquí desempeñan un lugar secundario y aparecen clasificadas de forma genérica en una única categoría—.

El calificativo de «complementarias» viene a poner el acento en el hecho de que este tipo de pruebas no facilitan de forma directa ni la toma de decisiones ni la clarificación del propio desarrollo personal, académico y profesional. Sin embargo, sí pueden proporcionar una información añadida a la que se recaba por los instrumentos anteriormente mencionados. Así, una persona que se está decidiendo por un campo profesional puede mostrar la necesidad o el interés por comprobar si reúne las capacidades necesarias para un buen desarrollo en esa profesión, o las características de personalidad más útiles para su desempeño. Además, es habitual que en los procesos de

selección de personal se utilicen este tipo de pruebas, por lo que es conveniente que la persona tenga un conocimiento previo y realista de la diferencia que puede existir entre su perfil de aptitudes y personalidad y el perfil exigido por el puesto de trabajo al que aspira.

Bajo estas aplicaciones, pruebas como las destinadas a evaluar las aptitudes, el rendimiento y los rasgos de personalidad tienen un gran potencial para el autoanálisis del sujeto joven y adulto, y van a contribuir a que éste defina y clarifique su autoconcepto. Sin embargo, hay que advertir del peligro que conlleva usar este tipo de pruebas para la clasificación y etiquetaje de las personas, o como un factor decisivo a la hora de predecir el rendimiento futuro en una profesión. Este uso de las pruebas de aptitud y rendimiento es el que anteriormente caracterizamos como de «portero», en tanto que los tests se utilizan para seleccionar a las personas y abrirles o no el acceso a ciertos campos de actividad profesional.

Dentro de este amplio grupo final de pruebas, en nuestro país disponemos de un considerable número de ellas que pueden utilizarse con fines orientadores. Así, de acuerdo con las variables que hemos introducido en este bloque, podemos citar como ejemplo una amplia gama de instrumentos. Para su selección, hemos manejado el criterio de que sean también aplicables a personas adultas y que los puedan administrar profesionales sin experiencia en diagnóstico clínico (esto no sucede, por ejemplo, con la mayor parte de los métodos proyectivos para la evaluación de la personalidad), además de presentar la versión más actual de la prueba en cuestión:

A) Pruebas para la **evaluación de la inteligencia, aptitudes y rendimiento.** Aquí incluimos:

— Pruebas para la evaluación del factor «g» (inteligencia general): *Tests de dominós D-48* y *D-70, Test de Factor G* de Catell y *Escalas de Matrices Progresivas* de Raven, entre las comercializadas por TEA, y el *Test de Inteligencia General* editado por Psymtec.

— Pruebas para la evaluación de las aptitudes intelectuales básicas (razonamiento abstracto, factor numérico, verbal, aptitud espacial, etc.): *Test de Aptitudes Diferenciales* DAT, *Test de Aptitudes Mentales Primarias* PMA, *Aptitudes Mentales Primarias Equivalentes* AMPE-F, etc., entre las que distribuye TEA. Asimismo, en Psymtec podemos encontrar la *Escala de Inteligencia Stanford-Binet,* el *Test de Inteligencia Verbal,* etc.

— Pruebas para la evaluación de aptitudes y conocimientos específicos requeridos en distintos oficios y profesiones. Del fondo editorial de TEA destacamos: *Test de Aptitudes Básicas para la Informática* (ABI), *Batería para la Actividad Comercial* (BAC), *Batería de Conductores* (BC), *Batería de Operarios* (BO), *Batería de Pruebas de Admisión* (BPA), *Batería de Subalternos* (BS), *Batería de Tareas Administrativas* (BTA), *Batería de Test de Oficios* DECATEST, *Test de Comprensión de Órdenes Escritas* (COE), *Test de Destreza en el manejo de*

Herramientas BENNET, *Test de Destreza con Pequeños Objetos* CRAWFORD, *Test de Instrucciones Complejas* IC, *Tests de Aptitudes Mecánicas* MacQUARRIE, *Test de Método y Orden* MO, *Prueba de Habilidades en la Negociación* NEGO, etc.

— Entre los tests de rendimiento escolar destaca el K-TEA o *Test de Mejora Educacional*, la *Batería R de Habilidades Cognitivas y de Conocimientos,* la *Batería de Valoración del Lenguaje,* el TALE o *Test de Análisis de la Lectoescritura,* comercializados por Psymtec. Las pruebas BOEHM o *Test de Conceptos Básicos,* BLOC o *Batería del Lenguaje Objetivo y Criterial,* la *Prueba de Comprensión Lectora,* EVOCA o *Estimación del Vocabulario,* etc. Muy conocidas son la *Batería Psicopedagógica* EOS y la *Batería Psicopedagógica* EVALUA, ambas comercializadas por EOS y disponibles para muchos niveles educativos.

B) Pruebas para la **evaluación de la personalidad.** Entre las pruebas generales, destacan: *Cuestionario Big Five* BFQ, *Cuestionario de Personalidad* CEP, *Inventario Psicológico de California* CPI, *Cuestionario de Personalidad Situacional* CPS, *Cuestionario de Personalidad* EPI, *Cuestionario de Personalidad Eysenck Revisado* EPQ-R, *Escalas de Adjetivos Interpersonales* IAS, *Inventario Tipológico de Myers-Briggs* MBTI, *Inventario de Personalidad* NEO PI-R, *Cuestionario Factorial de Personalidad* 16PF-5. Todas ellas son comercializadas por TEA.

Dada la cantidad de pruebas existentes para la evaluación de la personalidad y las capacidades intelectuales, no hemos indicado aquí todas las que existen, sino un conjunto representativo de ellas.

4. CONSTRUCCIÓN DE PRUEBAS OBJETIVAS

Según la línea a la que se vienen ajustando los capítulos anteriores, tras la revisión del concepto de test y sus características y tipos, tendríamos que hablar del proceso a seguir para la elaboración de estos instrumentos. Sin embargo, ésta no va a ser la forma de proceder en este capítulo. Y ello por una razón básica, todos los usuarios de tests no necesariamente van a ser constructores de este tipo de instrumentos. Esto no sucedía en la observación y la encuesta donde, en algunas ocasiones, se aplican instrumentos ya elaborados y, en otras, es el profesional del diagnóstico que va a usarlo quien tiene que elaborar un instrumento *ad-hoc* para una determinada situación.

Hemos insistido repetidas veces en que los tests son unos instrumentos muy exigentes en cuanto al tipo de medición que realizan. Ello implica que su proceso de elaboración no sólo sea lento y complejo, sino que exija unos conocimientos y destrezas que no se adquieren meramente por la realización de una titulación universitaria en Pedagogía o en Psicología. Estos conocimientos, por otra parte, se refieren a varias dimensiones: por un lado, son conocimientos que deben abarcar todo lo relativo al rasgo a medir (inteligencia, autoconcepto, intereses profesionales...) y las distintas teorías y enfoques respecto al mismo; por otro, son necesarios conocimientos avanzados sobre estadística y medición educativa (ANASTASI, 1992). Esto implica que, si bien un profesional del diagnóstico cualquiera podría elaborar un cuestionario, una entrevista, o una lista de control con bastante calidad, siempre y cuando siga un proceso riguroso, no necesariamente dispone de los conocimientos necesarios para construir un test. Es más, un profesional puede haber demostrado su cualificación elaborando una buena prueba de aptitud y no tener las competencias necesarias para elaborar otra de intereses profesionales, por desconocer, en este caso, la característica (los intereses profesionales) objeto de medición. Por esta razón, consideramos que explicar el proceso de construcción de un test escapa del objetivo de esta obra.

Sin embargo, la elaboración de pruebas objetivas de conocimiento sí que puede resultar un ámbito en el que, en ocasiones, cualquier profesional del diagnóstico tenga que participar. Es por ello que vamos a describir brevemente en este apartado el proceso de elaboración de pruebas objetivas de conocimiento y los aspectos que hay que tener en cuenta en dicho proceso. En cualquier caso, el proceso de elaboración de una buena prueba objetiva se parece bastante al de un test, por lo que aquí nos limitaremos a dar unas nociones generales que permitan que el profesional del diagnóstico sea capaz de adoptar una postura crítica, reflexiva e informada en la elaboración de este tipo de instrumentos. En nuestra exposición, recogemos parte de las tareas y recomendaciones propuestas por PÉREZ JUSTE (1991), las cuales resultan de gran claridad y utilidad para este fin.

Es necesario hacer constar que nos vamos a limitar a presentar el proceso de construcción de pruebas objetivas de elección múltiple; es decir, de aquellas en las que,

tras un enunciado, se presentan varias opciones de respuesta entre las cuales el sujeto debe elegir una, siendo sólo una de ellas correcta para la resolución del enunciado. La razón de esta elección se basa en que este tipo de pruebas son las más utilizadas para evaluar conocimientos. Aunque su proceso de elaboración comparte muchos aspectos y tareas en común con las pruebas cuyos ítems responden a otro formato (preguntas de verdadero-falso, de asociación de elementos, de completación de textos, etc.), cada uno de estos formatos presenta características propias, por lo que abordarlos todos resultaría demasiado extenso para esta obra, al tiempo que restaría claridad a la exposición.

Aunque a continuación explicitaremos los pasos a dar en la elaboración de pruebas objetivas, queremos transcribir aquí una cita de PÉREZ JUSTE (1989:121) con la que no podríamos estar más de acuerdo. Según este autor, «la fase de conceptualización de la prueba es, pues, la más difícil y delicada, además de la básica, ya que todo lo posterior viene radicalmente condicionado por ella». Por tanto, si esta fase se realiza con la suficiente prudencia, cuidado y meditación, estamos cumpliendo un requisito básico para conseguir que nuestra prueba tenga mejores garantías en cuanto a fiabilidad y validez.

4.1. Construcción de los planos de la prueba: objetivos, contenidos y tabla de especificaciones

MEHERNS y LEHMANN (1982:174) afirman que «las buenas pruebas no simplemente ocurren», sino que requieren una correcta y cuidadosa planificación que empieza con la clarificación de los contenidos y objetivos que se pretenden evaluar.

Por tanto, la primera tarea a realizar para elaborar una prueba objetiva de conocimientos consiste en especificar el material que se va a examinar. Es decir, se trata de delimitar cuáles son los contenidos que se quieren incluir en la prueba y explicitar su estructura en sub-unidades con entidad suficiente, aquellas sobre las que se organiza la enseñanza, lo que facilitará la representatividad de dicho contenido (PÉREZ JUSTE, 1989).

Suponiendo que deseáramos construir una prueba para examinar a estudiantes de Pedagogía en una asignatura sobre *Técnicas e instrumentos de diagnóstico*, tendríamos que especificar cuáles son los contenidos que se han trabajado y sobre los que queremos examinar a estos estudiantes. En el cuadro 2 presentamos un ejemplo de posibles contenidos. Se trata de elaborar un esquema en el que figuren todos los contenidos impartidos o, al menos, todos los contenidos que por el peso y relevancia que se ha dado durante la enseñanza, se considera necesario incluir en la prueba.

En opinión de PÉREZ JUSTE (1991), esta delimitación de los contenidos tiene varias utilidades:

a) No olvidar incluir ningún aspecto importante en el instrumento.

b) Incluir más elementos de aquellos apartados más amplios.

c) Ponderar cada apartado en función de su importancia intrínseca y relativa.

d) Si se pretende obtener puntuaciones o diagnósticos diferenciados, se podrá llegar a ponderaciones y estimaciones del nivel de los sujetos en las distintas pruebas.

Además de los contenidos, es conveniente especificar los objetivos de aprendizaje que se habían marcado para esos contenidos y hacia los cuales se ha orientado el proceso de enseñanza y aprendizaje. Siguiendo con nuestro ejemplo, con los contenidos mostrados en el cuadro 2, podemos haber perseguido que los estudiantes sólo los memoricen o que los comprendan, los apliquen, los relacionen... De esta forma, es evidente que con un mismo contenido pueden perseguirse objetivos de aprendizaje diferentes y esto debe quedar reflejado en la prueba.

CUADRO 2

EJEMPLO DE UN ESQUEMA DE CONTENIDOS A INCLUIR EN UNA PRUEBA OBJETIVA

Tema 1. Introducción al diagnóstico educativo

1. Desarrollo histórico del diagnóstico en educación.
2. Clarificación conceptual del diagnóstico educativo:
   - — Definiciones de diagnóstico.
   - — Diferencias y semejanzas del diagnóstico y la evaluación.
   - — Diferenciación con otros términos (evaluación psicopedagógica, evaluación de necesidades, evaluación de contextos...)
3. Funciones del diagnóstico educativo: apreciación, pronóstico, prevención...
4. Ámbitos de diagnóstico en educación: personal, socio-relacional y educativo.
5. Enfoques del diagnóstico educativo.
   - — Modelos tradicionales: atributo, dinámico y médico. Características generales.
   - — Modelos conductuales: conductual radical y conductual cognitivo. Características generales.
   - — Modelos cognitivistas: evaluación del potencial de aprendizaje y de los estilos cognitivos. Características generales.
6. Principios y modelos de intervención orientadora a tener en cuenta en los procesos diagnósticos:
   - — Principio de prevención, desarrollo e intervención social. Implicaciones cara al diagnóstico.
   - — Modelos de intervención por servicios y por programas. Implicaciones cara al diagnóstico.
7. Proceso de diagnóstico: características y etapas.
8. Informe diagnóstico: cuestiones a tener en cuenta y estructura.

Tema 2. Introducción a las técnicas e instrumentos de observación

1. Concepto y naturaleza de la observación en el diagnóstico educativo. Ventajas y limitaciones de los procedimientos observacionales.
2. Modalidades de observación: estructuración, participación y auto/hetero observación.
3. ...

En el cuadro 3, hemos sintetizado un ejemplo de posibles objetivos de aprendizaje que solemos usar en las pruebas objetivas mediante las cuales examinamos a nuestros sufridos estudiantes. Es necesario decir que estos objetivos son coherentes con la forma en que se trabaja la asignatura (que conlleva la realización de prácticas en grupo, el análisis crítico de instrumentos elaborados por los estudiantes, autoaplicación de

pruebas, etc.); si no fuera así, la prueba objetiva que estamos construyendo no tendría garantías de ser válida, ya que pretendería medir unos conocimientos que no se han podido adquirir durante el proceso de enseñanza/aprendizaje.

Visto lo anterior, parece claro que las pruebas objetivas no sólo sirven para medir conocimientos memorizados, sino que también resultan útiles para evaluar otros objetivos de aprendizaje más complejos.

CUADRO 3

EJEMPLO DE DEFINICIÓN DE LOS OBJETIVOS DE APRENDIZAJE A EXAMINAR EN UNA PRUEBA OBJETIVA

CONOCIMIENTO Y COMPRENSIÓN: Se refiere a que el estudiante ponga en práctica las siguientes actividades cognitivas con los contenidos de la asignatura: reconocer conceptos, comprender el significado de términos, comprender principios teóricos, memorizar términos concretos, identificar fases en un proceso, identificar principios en un diagnóstico, etc.

RELACIÓN: Se refiere a que el estudiante sea capaz de establecer relaciones entre los distintos contenidos del temario. Estas relaciones se aplican a contenidos de distintos temas, entre los cuales existe una separación temporal y física, y no tanto a relaciones que se dan ya como establecidas en la presentación del material de aprendizaje. Se trata, por tanto, de establecer relaciones entre contenidos de distintos temas y no entre los contenidos de un mismo tema o bloque.

APLICACIÓN: Incluye las actividades cognitivas necesarias para aplicar a situaciones o casos concretos los conceptos y principios incluidos en los contenidos. Se diferencia de los anteriores en tanto que se trata de resolver una cuestión relacionada con los contenidos en el marco concreto de un caso o situación que sirve de marco para la aplicación.

Una vez identificados estos dos importantes aspectos —cuáles son los contenidos y habilidades a examinar y poner en práctica por el estudiante en la prueba—, se procede a la confección de lo que se denomina como tabla de especificaciones. Esta tabla viene a ser el equivalente a los planos que utiliza un arquitecto para construir un edificio o al dibujo sobre el cual se confecciona un bordado. Tiene como finalidad servir de plano, de tal forma que, a partir de ella se decidirá sobre el tipo y número de elementos a incluir en la prueba.

Existen varias formas de confeccionar la tabla de especificaciones. Puede decirse que cada profesional utiliza el enfoque que se adapta mejor a su forma de trabajar. Algunas tablas de especificaciones consisten en un listado de temas, como el que recogíamos en el cuadro 2, en el que se indica cuántos ítems se van a incluir para cada tema (según la

relevancia que cada bloque puede tener en función de cómo se ha impartido la materia a examinar).

Otra forma de realizar la tabla de especificaciones consiste en un cuadro de doble entrada en el que se cruzan cada objetivo con cada contenido[18], de forma que el constructor de la prueba decide qué tipo de aprendizajes se van a relacionar con qué contenidos. Así, no es necesario que todos los contenidos se examinen de acuerdo con todos los objetivos, sino que en función de la enseñanza impartida, se pueden examinar ciertos objetivos con ciertos contenidos más que con otros.

Nosotros nos decantamos por una tabla de especificaciones en la que se crucen diversas informaciones:

— Por un lado, como en el último procedimiento indicado, hay que decidir qué objetivos se van a examinar para qué contenidos. La forma en que se ha impartido la materia condicionará esta decisión. Así, si no se han realizado prácticas de un determinado tema, es mejor que en sus contenidos sólo se examine la comprensión y la relación.

— Por otro, para que la tabla de especificaciones resulte útil para las tareas a realizar con posterioridad, es importante decidir sobre el número de preguntas que finalmente va a contener la prueba y, en función de esto, el número de preguntas que va a incluirse en cada bloque de contenidos y para cada objetivo.

— Asimismo, consideramos importante que en la tabla de especificaciones, el constructor también prevea el nivel de dificultad de los ítems y cómo se van a combinar los ítems de distintos niveles de dificultad con los distintos contenidos y objetivos a examinar. Sobre este último aspecto, no obstante, hay que hacer algunos comentarios adicionales.

El nivel de dificultad de los ítems es una propiedad que caracteriza a los reactivos en los tests y pruebas objetivas a la que aludiremos en el último apartado de este capítulo. En realidad, es un índice que se calcula de forma empírica como la proporción de sujetos que han contestado correctamente al ítem. Si un ítem es contestado correctamente por una gran mayoría de los sujetos a los que se examina, podemos decir que es un ítem fácil. Su dificultad será, por tanto, mayor o menor, según el número de sujetos que lo cumplimenten satisfactoriamente.

A la hora de construir por primera vez una prueba, lo normal es que su constructor no disponga de esta evidencia empírica sobre la dificultad de los ítems, puesto que los está creando en ese preciso momento, por lo que, cuando aquí nos hemos referido a incluir en la tabla de especificaciones el número de ítems con distintos grados de dificultad, hablamos de una estimación intuitiva por parte de quien elabora la prueba. Seguramente, la persona que ha impartido la asignatura (coincida o no con el constructor de la prueba), puede intuir cuando un determinado ítem puede resultar de dificultad media, baja o alta. Debe hacerse constar, en todo momento, que ésta es una aproximación exploratoria,

intuitiva, que no tiene fundamento empírico alguno, pero que puede resultar de utilidad para construir una buena prueba. Si el examinador que construye la tabla de especificaciones cuenta con un banco de ítems que ya han sido utilizados y estudiados, podrá tomar estas decisiones sobre la base más firme de la estadística, que sería lo deseable en todos los casos.

De cara a la construcción de la tabla de especificaciones, se trata de decidir cuántos ítems de dificultad media, alta y baja se van a incluir, y cómo se van a distribuir entre los distintos contenidos y objetivos de aprendizaje.

En el cuadro 4 hemos sintetizado y ejemplificado todas estas orientaciones. Para no hacer demasiado complejo el ejemplo, se ha preferido no indicar exhaustivamente todos los contenidos, sino sólo los más importantes, así como hacer la tabla para sólo dos temas (tema 1 y tema 2). En realidad, la tabla debe contener todos los contenidos a examinar.

Un rápido vistazo a la tabla construida nos permite observar algunas decisiones que ha tomado el constructor. En primer lugar, observamos que ha decidido que la prueba contendrá un total de 30 ítems. Asimismo, ha considerado que el tema 2 es más importante y se ha trabajado más que el 1 en clase, por lo que ha decidido que un mayor porcentaje de los ítems tendrán como contenido dicho tema. Otra decisión importante se refiere al número de ítems de cada nivel de dificultad —recordemos que ésta es una apreciación intuitiva—. Cuando veíamos la diferencia entre los tests normativos y criteriales, apuntábamos que la decisión sobre la distribución de ítems según su dificultad se toma de forma diferente según el tipo de prueba. En nuestro ejemplo, se ha considerado que es conveniente que predominen los ítems con dificultad media (por eso constituyen un total de 24 sobre 30).

También se ha tomado una decisión acerca del número de ítems que se van a utilizar para examinar cada uno de los tres grandes objetivos. Se observa así que han predominado los ítems que evalúan el conocimiento y comprensión de conceptos. Esta decisión debe estar basada en el proceso de enseñanza/aprendizaje que se ha seguido. En este caso concreto, se ha considerado que el tema 1 engloba fundamentalmente contenidos conceptuales y que las prácticas realizadas han sido pocas, por lo que se ha incluido un porcentaje menor de ítems de aplicación. Asimismo, con sólo dos temas (como sucede en nuestro ejemplo) resulta difícil establecer relaciones entre contenidos, por lo que es mejor que haya pocos ítems para este objetivo.

Además de estas cuatro «grandes» decisiones (número total de ítems, número total de ítems en cada tema, número total de ítems según nivel de dificultad y número total de ítems por objetivo), el constructor de la tabla de especificaciones de nuestro ejemplo ha tomado muchas otras. Por ejemplo, ha decidido que la mayor parte de los ítems de aplicación se van a incluir dentro del tema 2. Esta decisión posiblemente se fundamenta en el proceso de enseñanza seguido. Así, en ese tema posiblemente se han realizado más prácticas, por lo que el constructor ha decidido que el número de ítems de aplicación de

ciertos contenidos del tema sea mayor que en otros contenidos. Son muchas las decisiones como éstas que se han tomado y resultaría repetitivo analizarlas una a una. En cualquier caso, la idea principal que debe imperar en la toma de este tipo de decisiones es que sean coherentes con el planteamiento didáctico seguido, de tal forma que la prueba que finalmente se elabore contenga una muestra representativa de las tareas de aprendizaje a evaluar.

CUADRO 4

EJEMPLO DE TABLA DE ESPECIFICACIONES

Objetivos / Contenidos	CONOCIMIENTO Y COMPRENSIÓN	RELACIÓN	APLICACIÓN	Total ítems
TEMA 1	**7** / 1_b- 6_m-0_a	**2** / 0_b-1_m-1_a	**1** / 1_b-0_m-0_a	**10**
Concepto de diagnóstico	1_m			1
Funciones	1_m			1
Ámbitos	1_b			1
Enfoques	1_m	1_a		2
Principios y modelos de intervención	1_m			1
Proceso	1_m		1_b	2
Informe	1_m	1_m		2
TEMA 2	**13** / 0_b-13_m-0_a	**3** / 1_b-2_m-0_a	**4** / 0_b-2_m-2_a	**20**
Concepto de observación	2_m			2
Modalidades de observación	2_m			2
Preparación de la observación	2_m	1_b		3
Tipos de registros	2_m		1_m	3
Construcción de registros sistematizados	2_m	1_m	1_a	4
Preparación y aplicación de registros no sistematizados	2_m		1_m	3
Garantías científicas	1_m	1_m	1_a	3
TOTALES	**20** / 1_b-19_m-0_a	**5** / 1_b-3_m-1_a	**5** / 1_b-2_m-2_a	**30**

Total de ítems de la prueba según temas: 30 (33,3% del tema 1 y 66,7% del tema 2)
Total de ítems según dificultad: 30 (10% dificultad baja, 80% dificultad media y 10% dificultad alta)

Probablemente el constructor de la prueba tenga que dedicar bastante tiempo a la realización de las tres tareas descritas en este apartado (delimitación de contenidos y objetivos y elaboración de la tabla de especificaciones). Sin embargo, las decisiones importantes ya han sido tomadas, gracias a ellas. Podemos decir que se han construido

los planos; ahora sólo queda hacerlos realidad, llevando a la práctica las decisiones tomadas. Como plantean BROWN (1980) y MEHRENS y LEHMANN (1982), la tabla de especificaciones no sólo proporciona una base para construir reactivos o preguntas, sino que también sirve para determinar la validez de contenido de la prueba.

4.2. Formulación y selección de elementos

Ésta es una de las etapas más interesantes y creativas de la elaboración de pruebas objetivas aunque, a veces, puede requerir un considerable esfuerzo de imaginación. Básicamente consiste en formular ítems que sean indicativos de las tareas designadas en la tabla de especificaciones. Es decir, que representen la peculiar combinación de contenidos, objetivos y niveles de dificultad allí especificados, por lo que la formulación de los ítems es un momento clave en la construcción de pruebas. Más adelante profundizaremos en esta cuestión. Antes nos parece necesario explicitar dos supuestos de los que partimos a la hora de enfocar este apartado.

En primer lugar, hay que recordar que estamos explicando el proceso de construcción de pruebas en las que los ítems son de elección múltiple. Si no hubiéramos acotado el proceso en este sentido, llegados a este punto, la cuestión se complicaría algo más: habría que hablar de los distintos tipos de ítems, de sus características y de cómo se deben formular.

En segundo lugar, partimos de que la mejor forma de elaborar una prueba es formulando un número más amplio de elementos de los que realmente se van a usar, tal y como recomiendan la mayoría de los autores (BROWN, 1980; PÉREZ JUSTE, 1991; etc.). De esta forma, los elementos o ítems que finalmente se seleccionen para ser incluidos en la prueba serán los de más calidad y también los que mejor se adapten a las especificaciones de la tabla que elaboramos en la fase anterior.

Teniendo esto en cuenta, se trata ahora de acudir a nuestros conocimientos de la materia que impartimos, del proceso didáctico seguido y, sobre todo, a nuestra imaginación, de tal forma que seamos capaces de plasmar las decisiones tomadas en la tabla de especificaciones en un conjunto amplio de elementos entre los cuales seleccionaremos los mejores.

Utilizaremos la tabla de especificaciones del cuadro 4 para poner algunos ejemplos. Es necesario recordar que aquí partimos de una estimación intuitiva del nivel de dificultad de los ítems. Es decir, de una apreciación de qué enunciados pueden resultar más o menos difíciles a los estudiantes a los que va dirigida la prueba, apreciación que debe realizar la persona que ha impartido la materia, que es quien puede tener un juicio más informado al respecto.

Uno de los ítems que tenemos que formular, según nuestra tabla de especificaciones, ha de referirse a los contenidos denominados «principios y modelos de intervención» del

tema 1. Debe servir, asimismo, para evaluar el nivel de conocimiento y comprensión del concepto, además de tener un nivel de dificultad medio. Un ejemplo de ítem con estas características es:

EJEMPLO 1:

¿Cuál de las siguientes características es impropia del modelo de intervención por programas?

a) Atención al principio de intervención social.

b) Intervención de carácter reactivo.

c) Intervención determinada por las necesidades detectadas.

d) Carácter procesual de la intervención.

Se puede observar que lo que se pide es que el estudiante comprenda y conozca las características de la intervención por programas. En el ejemplo 2, se ha elegido el contenido «enfoques» del tema 1 para hacer un ítem de relación de dificultad más o menos alta.

EJEMPLO 2:

¿En cuál de los siguientes enfoques de diagnóstico es habitual el uso de sistemas categoriales de observación?

a) En los modelos propios de la perspectiva tradicional.

b) En los modelos propios de la perspectiva cognitivista.

c) En los modelos propios de la perspectiva conductual.

d) B y C son correctas.

En este ejemplo se pide al estudiante que relacione lo que sabe sobre los enfoques de diagnóstico y sus características propias con lo que sabe sobre los sistemas categoriales de observación. Son contenidos que pertenecen a dos temas distintos y a los que el material de aprendizaje no hace ninguna referencia expresa, por lo que la relación entre los contenidos la tiene que establecer el estudiante.

Un último ejemplo que vamos a presentar es el de un ítem que pretende comprobar si el estudiante es capaz de aplicar un contenido del tema 2 relativo al apartado de «construcción de registros sistematizados». Aquí la dificultad también es alta, más si cabe que en el ejemplo anterior.

EJEMPLO 3:

Un maestro cree que las niñas se comportan mejor que los niños. Si en la observación del comportamiento de sus alumnos/as, mediante una escala de estimación, suele otorgar valoraciones más altas a las niñas si las conductas son positivas y más altas a los niños

cuando las conductas son negativas, ¿qué efectos o errores están incidiendo en su observación?

a) Efecto de halo.

b) Error de indulgencia.

c) Error de severidad y efecto de halo.

d) Error de indulgencia y de severidad.

Este es un ejemplo en el que se aprecia que el estudiante tiene que aplicar lo que sabe sobre los errores y efectos que inciden en las escalas de estimación al análisis de una situación concreta.

Como se puede apreciar, la formulación de los elementos de la prueba requiere ciertas dosis de imaginación y creatividad, además de un conocimiento profundo de los contenidos a evaluar y de la forma en que se han impartido y desarrollado los mismos en el proceso didáctico seguido. El constructor de la prueba deberá dedicar a esta tarea bastante tiempo y volver a ella una y otra vez hasta que consiga reunir un número amplio de elementos para la prueba que quiere elaborar.

En cualquier caso, existen dos formas de proceder a la hora de formular los elementos. Una de ellas puede caracterizarse como *deductiva*. Es similar a la forma en que hemos procedido aquí con los ejemplos que hemos puesto. Se trata de identificar primero qué tipo de ítem hay que elaborar y después proceder a su formulación. Por ejemplo, miramos nuestra tabla de especificaciones y vemos que el primer elemento a formular debe ser del tema 1 (contenido «concepto de diagnóstico»), dirigido a evaluar conocimiento y comprensión y de una dificultad media; identificados estos parámetros, procedemos a formular un ítem que responda a los mismos.

Una segunda forma de proceder, que tiene algunas ventajas respecto a la primera, es más *inductiva*. En este caso y con el material de aprendizaje por delante (apuntes, libros u otros materiales que se han usado para impartir la materia), el constructor de la prueba se dispone a formular cuantos ítems se le ocurran, teniendo siempre presente el contenido y los objetivos que los ítems deben representar y la necesidad de que tengan diferentes niveles de dificultad. Una vez que ha elaborado un banco de ítems lo más amplio posible, se trata ahora de que identifique los parámetros de cada ítem: contenido que examina, objetivo que evalúa y nivel (intuitivo) de dificultad del ítem. Una ventaja de este segundo procedimiento es el mejor aprovechamiento de la creatividad del constructor de la prueba, ya que conforme inventa un ítem es posible que se le ocurran otros distintos dirigidos a los mismos u otros parámetros. Sin embargo, la ventaja más importante de este procedimiento inductivo es que, a través de él, resulta más fácil hacer una contrastación previa con otros colegas. Así, una vez elaborado el banco de ítems inicial, se puede pasar a otros profesionales que: a) dominan los contenidos, b) discriminan entre los objetivos por su conocimiento del proceso didáctico que se ha seguido en la materia y c) conocen a los estudiantes lo suficientemente bien como para

poder hacer una estimación intuitiva del grado de dificultad de los ítems. Estos profesionales van a clasificar los ítems, desde su perspectiva, en función de sus contenidos, objetivos y nivel de dificultad, lo cual puede favorecer la comparación y la consecución de una clasificación definitiva más exhaustiva y fidedigna.

No obstante todas las consideraciones realizadas en los párrafos precedentes, hay otros aspectos de gran importancia en la formulación de ítems. Determinadas cuestiones se deben tener presentes siempre que se formulan elementos para una prueba objetiva, ya que van a contribuir a la calidad final de los ítems formulados. Estas cuestiones pueden dividirse en dos grandes grupos: por un lado, las generales, que se refieren a todo tipo de elemento, sea o no de elección múltiple; y, por otro, las específicas de cada tipo de elemento (en nuestro caso, presentaremos aquí sólo las cuestiones específicas referidas a los ítems de elección múltiple).

Respecto a las cuestiones generales a tener en cuenta en la formulación de todo tipo de elementos, hemos revisado las aportaciones de diversos autores (BROWN, 1980; GARCÍA JIMÉNEZ, 1997; HALADYNA, 1994; MEHRENS y LEHMANN, 1982; PÉREZ JUSTE, 1989 y 1991; THORNDIKE y HAGEN, 1978) y las hemos sintetizado en la lista siguiente:

— El enunciado del elemento debe ofrecer toda la información necesaria, pero sólo la necesaria, para que los sujetos puedan dar o reconocer la respuesta.

— La formulación del enunciado debe ser clara. El vocabulario utilizado debe ser correcto y adecuado al nivel de los contenidos (y sujetos) que se van a examinar.

— El enunciado debe ser cuanto más breve mejor, de forma que figure en él sólo lo estrictamente necesario y funcional.

— El enunciado debe presentar corrección gramatical, ortográfica y sintáctica.

— Cada elemento debe referirse a un sólo contenido-objetivo.

— Asegurarse de que existe una respuesta correcta u óptima para cada ítem. Esto es especialmente evidente cuando en el contenido que se somete a examen existen diversas posturas posibles. En esos casos, el ítem debe indicar claramente cuál es la opinión o la autoridad que ha de emplearse como fundamento de la respuesta.

— Evitar hacer citas literales de los textos que se han usado para preparar el examen. Además del hecho de que las citas literales sacadas de su contexto pueden perder parte de su significado, esta forma de formular los enunciados no permite ir más allá del simple recuerdo o memorización.

— Los elementos deben redactarse en forma afirmativa, evitando las dobles negaciones. En ocasiones, un «no» en el enunciado puede pasar desapercibido, además de que puede originar que la pregunta sea entendida de forma incorrecta. Por ejemplo: El informe de diagnóstico no es un documento escrito: SÍ/NO ¿Si el sujeto responde SÍ, significa que «sí, no es un documento escrito», y si responde

NO, que «no, no lo es»? Como se puede apreciar, resulta difícil saber cómo hay que contestar a esta pregunta, independientemente de que se domine o no el contenido.

Entre las cuestiones específicas de los elementos de elección múltiple, nuevamente y teniendo en cuenta distintas aportaciones (BROWN, 1980; GARCÍA JIMÉNEZ, 1997; HALADYNA, 1994; MEHRENS y LEHMANN, 1982; PÉREZ JUSTE, 1989 y 1991; THORNDIKE y HAGEN, 1978), podemos señalar las siguientes:

— Presentar, junto a la respuesta correcta, buenas alternativas o distractores que puedan atraer la atención del sujeto como posible respuesta correcta. Estos distractores pueden obtenerse a partir de respuestas erróneas habituales o utilizando palabras o formulaciones técnicas.

— El número óptimo de opciones de respuesta oscila entre 3 y 5 (MEHRENS y LEHMAN, 1982), aunque otros autores (THORNDIKE y HAGEN, 1978) recomiendan el uso de 4 o 5 opciones. Teóricamente, a mayor número de distractores plausibles, mayor es la fiabilidad, pero en ocasiones resulta muy difícil encontrar varios distractores plausibles, al tiempo que cuantas más opciones de respuesta haya, mayor será el tiempo necesario para realizar la prueba.

— Evitar repetir palabras en las distintas opciones. La raíz del enunciado debe redactarse de forma que se incluyan en ella las palabras necesarias para evitar la repetición en las alternativas de respuesta.

— Proyectar el contenido del ítem hacia situaciones novedosas si se desea evaluar procesos mentales superiores. Los ítems de elección múltiple se prestan a la evaluación de ciertas capacidades (comprensión, establecimiento de relaciones, aplicación...), pero es necesario presentar al sujeto una situación novedosa para la que se requiera algo más que el recuerdo de lo que aprendió de memoria. Así, debe evitarse usar en las preguntas las mismas frases y ejemplos que se han utilizado durante el proceso de enseñanza.

— Las respuestas correctas deben situarse en un orden que se establezca al azar, de forma que el sujeto no se pueda percatar de la pauta que se utiliza y responda según la misma. En la redacción de cada ítem, en el caso de que las alternativas tengan entre sí un orden lógico, conviene ponerlas en ese orden. Por ejemplo, en orden ascendente o descendente si se trata de números, por orden alfabético si se trata de palabras aisladas...

— En relación con esto, el constructor de la prueba debe procurar que el porcentaje de respuesta correcta en cada opción sea semejante. Por ejemplo, si tenemos un ítem de elección múltiple que tiene cuatro opciones, A, B, C y D, la respuesta correcta debe estar en proporciones semejantes en las cuatro opciones, aunque tampoco se trata de que el número de Aes, Bes, Ces y Des sea exactamente el mismo, lo que produciría un efecto contraproducente. En ocasiones, los constructores de pruebas tienden a colocar las respuestas correctas en las opciones centrales (B y C), porque

les parece que así quedan mejor «escondidas». Es por ello que, en muchas pruebas se detecta que existe una sobreabundancia de respuestas B y C. Si el sujeto detecta esta pauta, ante una cuestión en la que dude, puede tender a responder una de las dos opciones centrales porque sabe que así tendrá más probabilidad de acertar.

— Las alternativas tienen que ser homogéneas en la forma y la estructura. Es muy importante que ninguna de las opciones ofrecidas —entre ellas, la correcta— pueda quedar eliminada por la falta de concordancia gramatical entre el enunciado y la opción de respuesta, o por cuestiones similares. Asimismo, tampoco la respuesta correcta debe ser seleccionada por presentar una forma diferente a las demás. Hay que evitar dar indicios poco relevantes sobre ella mediante aspectos como su longitud, la repetición de palabras clave, etc.

— Cada pregunta ha de ser independiente. Un reactivo no debe ayudar a responder a otro de la prueba.

A continuación, presentamos algunos ejemplos de ítems que recogen o no algunas de las cuestiones apuntadas, tanto las generales como las específicas.

EJEMPLO 4:

Los intervalos de un sistema de observación deben ser:

a) Cortos, ya que así resulta más fácil observar y registrar.

b) Cortos, ya que así aumenta la fiabilidad del sistema.

c) Largos, ya que así disminuye la reactividad de la conducta.

d) Cortos, ya que así se consigue mejor generalizabilidad.

EJEMPLO 5:

En los planteamientos iniciales del diagnóstico (corriente médico-psicológica):

a) No existía el diagnóstico.

b) La actividad diagnóstica estaba vinculada al contexto de aprendizaje.

c) El diagnóstico y la evaluación eran dos actividades esencialmente iguales.

d) Sólo el alumnado con problemas se beneficiaba del diagnóstico.

EJEMPLO 6:

Señala la respuesta incorrecta entre el siguiente conjunto de características de la intervención en el modelo por programas:

a) La intervención es procesual.

b) La intervención es remedial.

c) Se detectan necesidades.

d) La intervención es preventiva.

EJEMPLO 7:

Cuando el observador que utiliza una escala de estimación tiende a conceder puntuaciones bajas a todos los sujetos que observa, es posible que cometa:

a) El error de tendencia central.

b) El error de lenidad.

c) El error de indulgencia

d) El error de severidad.

En el ejemplo 4, observamos que la opción C es muy diferente a las demás. En las otras, la respuesta siempre es «cortos», sea por una razón o por otra. Sin embargo, en C la respuesta es «largos». Esta falta de isomorfismo entre la opción C y las demás puede llevar a que el sujeto la rechace o acepte no en función de lo que sabe, que sería en todo momento lo deseable, sino en función de características externas o de forma del ítem. Si otra opción hubiera empezado por «largos», el sujeto tendría que prestar mayor atención para discriminar su respuesta.

En el ejemplo 5, la opción A es absurda. Si el enunciado parte de «los planteamientos iniciales» del diagnóstico, evidentemente no puede decirse que «no existía el diagnóstico». Es una opción que puede descartarse sin necesidad de haber cursado la materia.

En el ejemplo 6, la opción C también presenta una diferencia evidente respecto a las demás. En todas las demás, el esquema sintáctico es: «la intervención» como sujeto de la frase, el verbo «es» y después un predicado constituido por un adjetivo, que aquí realiza la función de atributo. La estructura sintáctica de la frase C responde a un esquema totalmente distinto. Además de esto, mientras las otras tres opciones concuerdan semánticamente con el enunciado, el cual hace referencia explícita a las «características de la intervención», la frase C no concuerda con el enunciado de forma tan evidente.

Por último, en el ejemplo 7 se observa cómo se repite la expresión «el error de...» en todas las alternativas de respuesta. El pie de la pregunta podría haber sido formulado en otros términos para evitar esta repetición: «Cuando el observador que utiliza una escala de estimación tiende a conceder puntuaciones bajas a todos los sujetos que observa, es posible que cometa el error de...». Esta segunda forma de redactar el enunciado ahorrará tiempo de lectura al sujeto, además de que le ayudará a concentrar su atención en el problema.

Del ejemplo número 6 queremos resaltar la palabra «incorrecta». Esta palabra, y otras similares, son las que nos pueden permitir evitar el uso del «no» en los enunciados de los elementos. Así, en este caso, para no decir «señala la respuesta que no es correcta», se ha utilizado la expresión «incorrecta». Otras expresiones similares que pueden servir a este

fin son «inadecuada» por «no adecuada», «impropia» por «no propia», etc. También verbos como «evitar» o «eludir», etc. pueden servir a esta finalidad, sustituyendo a expresiones como «lo que no se debe hacer», «lo que no se hace», etc. El uso de adjetivos o expresiones contrarias es otra forma de evitar la negación. Por ejemplo, «hacer un informe es una tarea diferente...» por «hacer un informe no es una tarea igual a...», etc.

Como puede apreciarse, la formulación de elementos no es una tarea fácil, ya que son muchos los aspectos que hay que tener en cuenta, tanto en lo relativo al contenido del ítem (nivel de dificultad y objetivos y contenidos que evalúa), como en lo relativo a su forma. Por eso siempre resulta muy apropiado que en su elaboración participen varias personas. Esto contribuirá a que los ítems estén más depurados y mejor clasificados, además de que la creatividad individual se puede ver favorecida por las aportaciones de los demás.

4.3. Ordenación de los elementos y preparación del formato de la prueba

Una vez decidido cuáles son los ítems que se van a incluir en la prueba, se procede a preparar el formato de la misma. Para facilitar su corrección, es conveniente diferenciar dos partes en ella: por un lado, el cuadernillo en el que figuran las preguntas o ítems y, por otro, una hoja ya preparada para que el sujeto consigne sus respuestas. Esta forma de presentación puede ser también útil para abaratar los costes de reproducción. Así, el cuadernillo de preguntas —que suele contener más folios— es reciclable: lo pueden usar varios sujetos. Para ello, en las instrucciones se debe especificar a los mismos que no deben anotar nada en los cuadernillos de preguntas.

Además del nombre y algunos datos complementarios[19] del sujeto, la prueba debe contener las instrucciones necesarias para su cumplimentación. Estas instrucciones, figuren o no por escrito en la prueba, deben ser explicadas en voz alta en el momento de su aplicación, a fin de asegurarnos de que todos los sujetos han comprendido los términos en que deben actuar (tiempo total del ejercicio, lugar y forma de consignar la respuesta, etc.).

Una de las cuestiones más importantes en la preparación del formato de la prueba se refiere a cómo se deben ordenar los elementos dentro de ella. De esta forma, se pueden utilizar diferentes criterios de ordenación (PÉREZ JUSTE, 1991): según el orden lógico o cronológico con que se ha impartido el temario, según el nivel de dificultad de los ítems, etc. Cada una de estas opciones presenta ventajas e inconvenientes.

Cuando se hace una ordenación en función de la secuencia con que se ha impartido la materia y todas las preguntas de un mismo tema aparecen de forma consecutiva, se facilita el recuerdo del sujeto. Éste tiene que situarse en un determinado tema para responder y así le resulta más fácil comprender los enunciados y contestar a ellos.

La ordenación según la dificultad de los ítems presenta la ventaja de que introduce al sujeto progresivamente en el examen. El que los ítems más fáciles estén al principio le anima a seguir adelante, mientras que en el caso contrario, encontrar uno o varios ítems difíciles al principio puede condicionar su ejecución en el resto de la prueba.

Sin embargo, en ocasiones es preferible una ordenación al azar. Esto sucede, por ejemplo, cuando hacemos varias versiones de la prueba (modelo A, B...) para evitar que los sujetos puedan copiar. En este caso, una ordenación lógica, cronológica o por niveles de dificultad de los ítems puede favorecer a un grupo de sujetos respecto a otros. Si la versión A se confecciona mediante este criterio, y las siguientes se forman a partir de la versión A, puede resultar que la A, aun teniendo los mismos ítems que las demás, facilite una mejor ejecución por parte de los sujetos, por la peculiar ordenación de sus elementos, que no es la misma que en las demás. De todas formas, si se hace una ordenación al azar de los ítems, como aconsejan muchos autores (MEHRENS y LEHMANN, 1982; PÉREZ JUSTE, 1991), se debe procurar que en cada versión de la prueba no coincidan varios elementos difíciles al principio, a fin de que no causen desánimo en los sujetos.

En lo que respecta al tiempo que puede durar la prueba, éste dependerá de muchos factores: número de preguntas, dificultad de las mismas, preferencias del examinador, etc. En general, el tiempo total concedido no debe ser amplio, sino el justo y necesario. Hemos de recordar que, a diferencia de las pruebas de ensayo, en este tipo de pruebas se solicita una respuesta casi espontánea. El sujeto no sólo debe reconocer la respuesta correcta entre las incorrectas, sino que además debe dominar los contenidos lo suficientemente bien como para tener las ideas claras y responder correctamente sin necesidad de emplear mucho tiempo en ello.

La respuesta a dar no se trata de una «elaboración», sino de una «elección» y para ello el tiempo necesario es más limitado. Aunque, como plantea BROWN (1980), los límites de tiempo se deben establecer de modo que todos o casi todos los alumnos puedan concluir la prueba. Por otra parte, y esto es ya algo que nos atañe a los docentes, cuanto más tiempo «de sobra» tienen algunos estudiantes, más posible es que su mirada comience a vagabundear sospechosamente hacia el ejercicio del compañero.

En general, y aunque ya hemos dicho que depende de muchos factores, el tiempo que consideramos óptimo para una prueba objetiva está entre 1 y 1,5[20] minutos por cada ítem. Otra cosa será el que proporcionemos a los sujetos un mínimo entrenamiento sobre la forma en que deben responder, enseñándoles a desarrollar estrategias apropiadas para superar las pruebas objetivas, pero eso cae ya dentro del terreno de la enseñanza y la didáctica, en el que este libro no pretende entrar.

4.4. Valoración y puntuación

La valoración de los elementos de una prueba objetiva puede realizarse conforme a varios criterios. Uno de ellos, el más simple, es asignar un punto por cada respuesta correcta. En otras ocasiones, se puede plantear la posibilidad de dar más de un punto a la respuesta correcta en los ítems de dificultad más elevada. En estos casos el nivel de dificultad del ítem debe estar basado en datos de aplicaciones previas y no en el juicio subjetivo del constructor de la prueba.

De todas formas, en las pruebas con ítems de elección múltiple es necesario corregir el efecto del azar. Así, cuando ante un enunciado se presentan cuatro alternativas posibles, en un 25% de las ocasiones es posible que el alumno acierte por azar, independientemente de lo que sepa de la materia objeto de examen. Por ello, es habitual emplear la fórmula de corrección que figura en el cuadro 5.

No obstante, si la prueba elaborada se divide, a su vez, en partes (por ejemplo, primer y segundo parcial, o bloque de temas de observación, encuesta...), es conveniente obtener una puntuación diferenciada para cada una de las sub-pruebas. De esta forma, es posible precisar mejor el tipo y nivel de aprendizaje de cada estudiante que es lo que, al fin y al cabo, se persigue.

CUADRO 5

FÓRMULA PARA LA VALORACIÓN DE PRUEBAS DE ELECCIÓN MÚLTIPLE

$$\text{Puntuación} = \text{Aciertos} - \frac{\text{Errores}}{N - 1}$$

Donde N= número de alternativas de cada ítem (incluida la correcta)

5. APLICACIÓN DE TESTS: PRECAUCIONES Y COMPETENCIAS

Dado que los tests plantean importantes requisitos para su aplicación y su interpretación, no son instrumentos que puedan utilizarse sin demostrar unos mínimos formativos que garanticen su adecuado uso. Por ello, las editoriales que los comercializan utilizan una clasificación en tres niveles, en función de la especialización del instrumento. De esta forma, sólo se venden a los profesionales que demuestren tener los conocimientos necesarios respecto al nivel al que pertenece cada test. Estos niveles son:

— *Nivel A*. En él están incluidos todos aquellos tests que pueden aplicarse, puntuarse e interpretarse adecuadamente con la mera ayuda del manual del test, como es el caso de los tests de rendimiento.

— *Nivel B*. En este grupo se requiere un cierto conocimiento técnico sobre construcción y uso de tests, así como sobre los aspectos educativos o psicológicos subyacentes en la prueba. Muchos tests de aptitudes y cuestionarios de personalidad se hayan incluidos en este grupo.

— *Nivel C.* Requieren, además de un conocimiento profundo de los tests y de las teorías subyacentes, una experiencia supervisada con los instrumentos. Los tests proyectivos son un ejemplo de esta categoría.

Dejando a un lado los tests clasificados en el nivel C, que requieren una formación mucho más especializada que la que se adquiere simplemente en función de la titulación que se cursa en la Universidad, vamos a presentar algunas reflexiones sobre la forma en que deben usarse los tests en la práctica profesional. El lector comprobará que algunas de las cuestiones que vamos a mencionar aparecen, bien de forma explícita, bien de forma implícita, en la propuesta de código deontológico que hemos adjuntado en el anexo 2 de este libro.

Una de las primeras reflexiones a hacer, en lo relativo al uso de tests, se refiere a la *selección del instrumento* concreto con el que vamos a trabajar. Evidentemente esto dependerá de los objetivos del diagnóstico y también de la disponibilidad de instrumentos en el mercado. En ocasiones, ante ciertos rasgos, nos encontramos que no tenemos muchas pruebas entre las que elegir. Sin embargo, y como se ha demostrado en el mini-vademécum presentado, es habitual que en muchos casos, dispongamos de varios instrumentos que permiten medir los mismos rasgos o características. Se impone, pues, la necesidad de decidir cuál es el más adecuado entre ellos.

A la hora de seleccionar la prueba más adecuada para cada situación concreta, el criterio más importante que debemos tener en cuenta es su calidad. La calidad de un test, como se verá más adelante, dependerá de que la medida que aporte sea precisa y fiable (fiabilidad) y de su capacidad de ofrecer la información más válida de cara al fin para el

que está construido e, indirectamente, para el fin de nuestro diagnóstico (validez). Cuestiones como el coste del instrumento y el tiempo que requiere su administración e interpretación deberían estar en un segundo plano, y nunca en el primero.

Es por esta razón que entendemos que todo usuario de tests debe tener unos conocimientos mínimos relativos a la medida en educación. No se trata tanto de que los usuarios de tests tengan que ser expertos en validación de instrumentos de medida, sino de que sean capaces de adoptar una postura crítica ante la medida que proporciona un test determinado. Para ello es fundamental saber interpretar adecuadamente la información que el manual del test proporciona sobre los estudios de validez y fiabilidad realizados sobre el instrumento. Sólo de esta forma es posible tener una idea de las posibles limitaciones que el test puede presentar.

En ocasiones, se usan los tests sin ni siquiera haber leído la sección del manual en la que se detallan los estudios realizados para estimar las propiedades psicométricas del instrumento. Esto es una muestra de un uso poco crítico, informado y reflexivo de los tests, práctica que un buen profesional debe desechar.

Por otra parte, en apartados anteriores, hemos insistido en que los tests proporcionan una medida tipificada del comportamiento humano. Esto implica que es necesario *respetar escrupulosamente las condiciones* de aplicación, corrección e interpretación estipuladas en la guía técnica del test. Nunca se debe aplicar un test sin haber leído su contenido y el contenido del manual y, por supuesto, sin tener claras las condiciones en que se debe aplicar y el tipo de instrucciones que se deben ofrecer a los sujetos. En nuestra opinión, es incluso deseable que el profesional que va a usar por primera vez una prueba, comience autoaplicándosela. De esta forma, conocerá de primera mano las dificultades que los sujetos pueden tener a la hora de realizar las tareas, o simplemente, en el manejo del instrumento (cómo contestar en la hoja de respuesta, dónde leer las cuestiones, etc.).

Resulta imposible indicar aquí las condiciones concretas en que cada test se debe aplicar, corregir e interpretar, puesto que cada prueba tiene unas instrucciones y características propias. La única regla aplicable a todos los casos es seguir estrictamente lo especificado en el manual. En caso contrario, violamos los supuestos que sustentan la medida que vamos a realizar. Así, cuando a un sujeto se le ponen más ejemplos de los especificados en el manual, o se le dan más explicaciones de las que allí aparecen, estamos ofreciéndole una ayuda «extra» que los sujetos del grupo normativo no tuvieron. Después no resultará lícito comparar los niveles de nuestro sujeto con los del grupo normativo, puesto que las condiciones han variado de uno a otro.

Esto que indicamos en cuanto al respeto escrupuloso de las condiciones del test ha de extenderse a un planteamiento más general, que no se refiere ya a cada aplicación concreta de la prueba. Una de las prescripciones que aparecen sistemáticamente en los códigos deontológicos de las profesiones que trabajan con tests se refiere a la custodia de este tipo de instrumentos de evaluación. Bajo ningún concepto se pueden dar a conocer

las plantillas de respuestas o entrenar a los sujetos en la cumplimentación de tests. Cuando se hace esto se está invalidando automáticamente la medida que proporciona el instrumento. Es obvio que si un sujeto se sabe de memoria las respuestas a un test de inteligencia, ese test ha perdido en ese sujeto su capacidad de medir el rasgo. Este tipo de prácticas es de la máxima gravedad. En primer lugar, por las consecuencias derivadas respecto al sujeto, le estamos engañando si interpretamos que su nivel de inteligencia según el test es muy alto; en segundo lugar, el proceso de construcción de un test es tan complejo y arduo, que resulta aberrante deshacer en una sesión algo que ha costado años conseguir.

Un último grupo de advertencias a considerar en el uso de tests se refiere a la *interpretación de las puntuaciones*, tema que suele ser también de gran relevancia en los códigos deontológicos al uso. Como ya hemos apuntado, ésta debe realizarse de acuerdo con las instrucciones especificadas en el manual. Asimismo, se requieren ciertos conocimientos de estadística básica para hacer una correcta interpretación (de un percentil, de una puntuación típica, de una puntuación típica derivada, etc). Además, es preciso saber interpretar bien los rasgos que evalúa el test. Esto exige un conocimiento profundo, no sólo de los factores tal y como vienen definidos en el manual, sino también como fruto de lecturas e indagaciones previas. Cualquier profesional no puede interpretar las puntuaciones de un test de personalidad si no conoce bien el constructo que está estudiando y sus diversos rasgos y facetas.

Ajustar la interpretación de las puntuaciones a la definición del rasgo que da el manual del test es un requisito muy importante. En ocasiones, los estudiantes que aplican pruebas por primera vez cometen errores graves al ir más allá de lo que permiten los datos. Podemos poner por ejemplo el caso de un test que mide diversas facetas del autoconcepto. A la hora de interpretar el autoconcepto académico de una persona, podemos decir que «se considera un buen estudiante» o «que tiene de sí mismo una buena imagen de su actividad escolar». Nunca podremos decir que «es un buen estudiante», porque en ningún momento el test nos ha informado de esa cuestión. En el apartado del capítulo 1 dedicado al informe diagnóstico, se ofrecen algunas consideraciones que permiten profundizar en esta idea.

Para finalizar con este apartado, nos parece interesante transcribir aquí las competencias mínimas que han de observarse para un uso adecuado de los tests, las cuales resumen muchas de las reflexiones que aquí hemos realizado. Dichas competencias han sido formuladas por MORELAND et al. (1995) tras estudiar el nivel de formación de los usuarios de tests en el contexto estadounidense, y se refieren a los siguientes aspectos:

— Evitar errores al establecer las puntuaciones y registrar los datos.

— Abstenerse de etiquetar a las personas con términos despectivos basándose en sus puntuaciones en los tests.

— Mantener las plantillas de corrección y otros materiales en lugar seguro.

— Asegurarse de que todos los sujetos siguen las instrucciones correctamente.

— Utilizar instalaciones apropiadas para la aplicación de tests.

— Abstenerse de entrenar a los sujetos en los ítems del test, lo que conllevaría sesgo en los resultados.

— Estar dispuesto a interpretar los resultados y a dar consejo a los sujetos en sesiones diseñadas para ello.

— No hacer fotocopias de materiales originales.

— Abstenerse de utilizar hojas de respuesta hechas por uno mismo, que pueden conducir a errores cuando se aplica la correspondiente plantilla.

— Establecer una buena relación con los sujetos para que obtengan las puntuaciones que realmente les corresponden.

— Abstenerse de responder a las preguntas de los sujetos con más detalle del especificado por el manual del test.

— No asumir que una norma o baremo para una determinada situación o grupo de sujetos puede aplicarse a otra situación o grupo de sujetos distintos automáticamente.

6. INTERPRETACIÓN DE LAS PUNTUACIONES DE LOS TESTS

Tras la aplicación de un test, es habitual asignar valores numéricos a las respuestas de los sujetos, según una pauta concreta que dependerá de las normas de corrección y puntaje de cada prueba. Así, ante cada respuesta correcta[21] o que se asocie a un patrón dado, se asignará un determinado valor numérico. De tal forma que la suma de las puntuaciones obtenidas en el test completo y/o en las partes de que se compone (sub-tests), constituirán las puntuaciones directas de los sujetos en el rasgo o rasgos que mide el test.

Sin embargo, esta puntuación directa nos dice poco o nada respecto al nivel o estado del sujeto respecto a la característica objeto de evaluación. ¿Qué significado puede tener un 20 en una prueba de rendimiento en ciencias sociales? Puede resultar una puntuación alta o baja según el número de ítems que contenga la prueba. No será lo mismo tener un 20 sobre 90 ítems que tener un 20 sobre 30. Asimismo, también dependerá de lo fácil o difícil que esos ítems puedan resultar. Un sujeto que obtiene 20 en nuestro ejemplo, no necesariamente va a obtener la misma puntuación en una prueba similar (en cuanto a contenidos y número de ítems) si las preguntas de la segunda resultan más difíciles.

Es por ello que se impone la necesidad de tener una referencia que permita la comparación de la puntuación del sujeto de forma que podamos interpretarla. Esta comparación se puede realizar de acuerdo con diversos patrones. Se puede, por ejemplo, valorar la puntuación según ésta denote el alcance o no de unas metas u objetivos, en cuyo caso, estaríamos haciendo una interpretación con base en criterios. O se puede también comparar la puntuación de un sujeto con la obtenida por otros, determinando la posición relativa de dicho sujeto respecto a los demás. Esto es una interpretación con base en normas.

No nos vamos a ocupar aquí de la interpretación con referencia a criterios. En primer lugar, porque, como ya se ha indicado, todavía son escasos los tests que se basan en esta forma de interpretación de los resultados. En segundo lugar, la interpretación criterial dependerá de las pautas y criterios de interpretación propios de cada prueba concreta, por lo que poco puede aportarse al respecto.

En la interpretación normativa, en cambio, sí existen unos patrones comunes a la hora de proceder a la interpretación de los resultados de un test. En todos los casos, el procedimiento que se aplica es siempre el mismo: la puntuación directa obtenida por un sujeto es transformada en un tipo de medida que podemos denominar como *relativa*, siendo ésta la que permite la comparación con el grupo normativo. Vamos a explicitar mejor esta idea.

Cuando se construye un test normativo, además de los estudios empíricos previos que aporten evidencia sobre las propiedades psicométricas del instrumento construido

(fiabilidad y validez), se realiza un estudio de los resultados en el test de una muestra representativa de la población a la que la prueba va dirigida. Este estudio permite conocer valores tan básicos como cuál es la media de la muestra, la desviación típica de las puntuaciones, etc. A partir de esta información resulta posible elaborar unas tablas de transformación de puntuaciones (por ejemplo, conociendo la media y la desviación típica de un conjunto de puntuaciones, podemos transformar cualquier puntuación en una puntuación típica o *z*). Estas tablas se denominan como *baremos del test* y son imprescindibles para la interpretación normativa. Un test de referencia a normas no está completo sin los baremos, por muy cuidadosa que haya sido su elaboración y por muy positivas que sean sus evidencias de validez y fiabilidad.

Estos baremos deben ser lo más completos y específicos posible, de tal forma que no nos permitan sólo comparar al sujeto con todos los de la población, sino también, y sobre todo, con aquellos sub-grupos poblacionales que reúnen condiciones y características similares a las de nuestro sujeto. Veamos un ejemplo de esto: si estuviéramos utilizando una prueba de rendimiento en matemáticas, que es aplicable a todos los alumnos que cursan la educación primaria, sería poco adecuado comparar la puntuación directa de un sujeto que cursa 3º de primaria con la media y la desviación típica de los alumnos de 1º a 6º. Lo más apropiado sería comparar a nuestro alumno sólo con los de 3º que, al fin y al cabo, han recibido una enseñanza similar según los programas educativos por niveles y ciclos. Es más, sería deseable incluso que la puntuación del alumno del ejemplo se comparara sólo con la de los alumnos de su mismo sexo, especialmente si hay indicios significativos de que los chicos y las chicas presentan un rendimiento diferente en matemáticas.

En ocasiones, cuando los baremos de los tests no son completos, o resultan demasiado inespecíficos (en tanto no se han analizado por separado diferentes grupos poblaciones cuyos valores en el test podrían diferir), la interpretación que podemos hacer de las puntuaciones obtenidas a través de la prueba deja mucho que desear. Es por ello que un buen test debe ser lo más completo posible en sus baremos, además de indicar por qué no contiene baremos diferenciales para características que pudieran ser relevantes. Esto último se puede explicar también con nuestro ejemplo: si se ha realizado un contraste de las diferencias de medias en el rendimiento en matemáticas de los chicos y las chicas de 3º de primaria, y se ha comprobado que no existen diferencias significativas en las puntuaciones de ambos sexos, entonces un baremo conjunto (para chicas y chicos) no sería necesario, pero el test debe explicar y justificar a sus usuarios el por qué de esa decisión.

Volvamos ahora al procedimiento de interpretación normativa al que antes aludíamos. Decíamos que consistía en transformar la puntuación directa obtenida por el sujeto en una *medida relativa*. Sabiendo ya que la transformación se realiza mediante las tablas de baremos, vamos a explicar qué es esta medida relativa y cuántos tipos existen. De forma intuitiva, podemos decir que una medida relativa nos informa de la posición que ocupa un sujeto respecto de un grupo. Por ejemplo, si ordenamos una serie de puntuaciones de

la menor a la mayor, podremos saber qué rango u orden ocupa una determinada puntuación respecto a las demás (la primera, la decimosexta, etc.). Según ANASTASI (1982), la medida relativa o puntuación transformada a la que estamos haciendo referencia sirve a dos finalidades: por un lado, nos indica la situación relativa del sujeto respecto al grupo normativo, de forma que se puede valorar el nivel de ejecución de dicho sujeto en la prueba; por otro, la puntuación transformada nos permite comparar directamente el nivel de ejecución de un sujeto en dos o más tests distintos. Así, si aplicamos dos pruebas de inteligencia al mismo sujeto, la comparación de sus puntuaciones en ambas no se puede hacer sobre la puntuación directa, ya que ambos tests pueden tener diferentes números de ítems, diferentes tipos de tareas, diferentes niveles de dificultad en la tarea, etc. Para proceder a este tipo de comparación intra-individual, también es necesaria la transformación de la puntuación directa en una medida relativa.

Teniendo esto en cuenta, existen cinco tipos básicos de puntuaciones transformadas que suelen utilizarse para realizar la interpretación de los resultados de un test. Estos tipos son: las puntuaciones de nivel, las de edad, los percentiles, las puntuaciones típicas y las puntuaciones típicas derivadas (las dos últimas pueden considerarse también como un único grupo, dadas las similitudes que comparten). Aunque explicaremos en qué consiste cada una de ellas, nos centraremos más en las tres últimas que, además de ser las más utilizadas, pueden aplicarse a cualquier tipo de test.

Las **puntuaciones de nivel** tienen mayor aplicación y sentido en los tests de rendimiento. Si un test de rendimiento se dirige a una o varias etapas educativas, se calcularán las puntuaciones medias y las desviaciones típicas para cada grupo de alumnos de los distintos niveles o cursos educativos de que se compone la etapa (por ejemplo, las medias y desviaciones típicas de los estudiantes de 1º, de 2º, de 3º... de Primaria). La puntuación directa de cada sujeto se compara con la media de cada nivel y allí donde no existan diferencias significativas entre la media del sujeto y la del nivel, se asume como el nivel de rendimiento del sujeto (ANASTASI, 1982; MEHRENS y LEHMANN, 1982). Así, diremos que el sujeto X tiene un nivel en la materia Y de 6º de Primaria (que puede coincidir o no con el nivel de estudios que cursa).

Un procedimiento similar es el que se realiza cuando se utilizan **puntuaciones de edad.** Este segundo tipo de puntuaciones se ha usado para la inteligencia y las aptitudes intelectuales. En función de ellas se puede conocer si un determinado sujeto tiene una *edad mental* diferente o no a su edad cronológica (por encima, por debajo o igual a ésta). El concepto de edad mental alude al nivel de desarrollo intelectual del sujeto según sus resultados en el test y comparado con otros sujetos. Según ANASTASI (1982), el procedimiento de cálculo puede ser exactamente igual a la comparación de medias descrita para las puntuaciones de nivel, o también otro diferente que consiste en agrupar las pruebas por edad, como sucede en la escala de Binet-Simon, aunque algunos autores (BROWN, 1980; MEHRENS y LEHMANN, 1982; etc.) sólo hacen alusión a este último procedimiento como medio de estimación de la edad mental. De esta segunda forma, la

edad mental del sujeto será aquella correspondiente al nivel de las pruebas más alto que el sujeto puede ejecutar correctamente. Por poner un ejemplo, si un niño de 8 años completa satisfactoriamente la prueba para 10 años pero no la de 11, se infiere que su edad mental es de 10 años. En realidad, la edad máxima en la cual y por debajo de la cual un sujeto responde satisfactoriamente a todos los tests es la *edad basal*, a la que se añaden *bonificaciones parciales* en meses para todos los ítems que responda bien de tests de niveles de edad superiores.

La edad mental (EM), no obstante, tampoco ofrece mucha información, puesto que dependerá de su relación con la edad cronológica. Es por ello que habitualmente se usa el *cociente intelectual* (CI), que consiste en dividir la edad mental y la cronológica y multiplicar el resultado por 100. De esta forma, los sujetos con edad mental igual o similar a su edad cronológica tenderán a tener un cociente intelectual de 100, o cercano a esta puntuación. Aquellos que tengan una edad mental superior a su edad cronológica, obtendrán CI superior a 100, y los que tengan una edad mental inferior a su edad cronológica, tendrán un CI inferior a 100. Se obtiene así una medida fácilmente interpretable del nivel de desarrollo intelectual de una persona.

Sin embargo, las puntuaciones CI tienen importantes desventajas o limitaciones. La primera de ellas es que sólo se pueden aplicar a las pruebas de inteligencia, en las que tiene sentido hablar de «desarrollo intelectual» y de «edad mental». En segundo lugar, este tipo de puntuaciones pierde su capacidad discriminativa en la adultez, en la que la evolución de la inteligencia se produce a un ritmo más lento y, a todas luces, diferente respecto a la infancia. Además, no todas las pruebas satisfacen la condición de variabilidad uniforme del CI o de variabilidad progresiva de la EM —las desviaciones típicas son proporcionalmente iguales para las distintas edades— (ANASTASI, 1982), con lo que puede producirse un importante error en la interpretación. En las últimas décadas, para solventar este problema, se ha preferido usar un tipo de puntuación típica derivada denominada *cocientes de desviación CI* o *cocientes de desviación de inteligencia* (THORNDIKE y HAGEN ,1978).

Es importante mencionar que en el terreno educativo se han utilizado cocientes similares al de CI (PÉREZ JUSTE, 1991), como el *cociente de instrucción* (resultado de dividir la edad de instrucción por la edad cronológica y multiplicar por 100), y el *cociente de rendimiento intelectual* (resultado de dividir la edad de instrucción por la edad mental y multiplicar por 100).

Dada su aplicabilidad exclusiva en unos determinados tipos de tests, además de otras limitaciones, las puntuaciones de nivel y de edad no se usan tanto como los percentiles, las puntuaciones típicas y las puntuaciones típicas derivadas que pasamos ahora a analizar con más detenimiento[22].

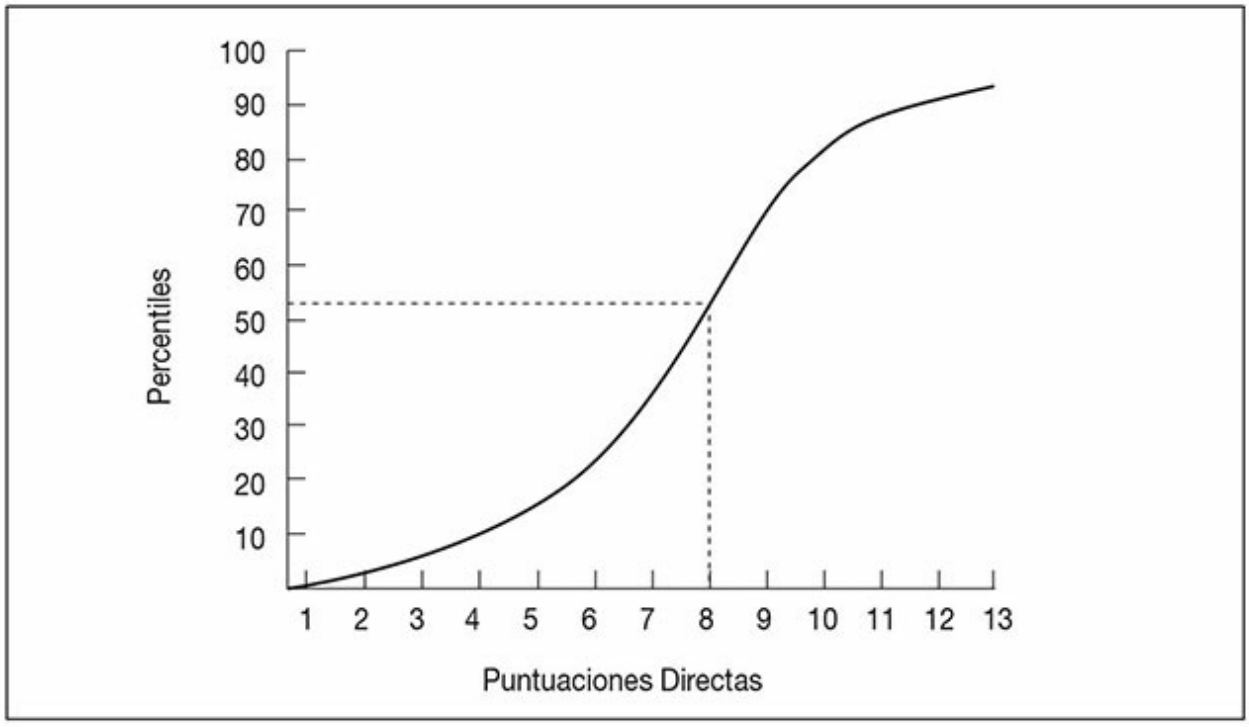
6.1. Puntuaciones percentiles

Las *puntuaciones percentiles o centiles* (P_n) constituyen una de las formas más sencillas y fáciles para interpretar los resultados de un test. Para su cálculo se ordenan las puntuaciones de un grupo de menor a mayor, de forma que se pueda determinar el porcentaje de sujetos que hay por debajo de una puntuación dada. Así, si decimos que una puntuación directa equivale al percentil 50 (P_{50}), estamos indicando que deja por debajo de sí misma al 50% de la población.

Los centiles suelen representarse gráficamente mediante la *ojiva de Galton*. En el eje de abscisas se sitúan las puntuaciones directas, mientras que en el de ordenadas se representa el porcentaje acumulado de casos que figuran por debajo de cada puntuación directa. En la figura 3 mostramos una representación gráfica de algunos percentiles.

FIGURA 3

OJIVA DE GALTON CON REPRESENTACIÓN DE ALGUNOS PERCENTILES



Las puntuaciones percentiles presentan un conjunto amplio de ventajas. Entre las más significativas podemos resaltar la facilidad que reviste tanto su cálculo como su interpretación. Así, conociendo el porcentaje de casos que quedan por debajo de una puntuación directa dada, podemos hacernos una idea bastante aproximada del lugar que ocupa el sujeto respecto al grupo normativo en función de su ejecución en el test. Asimismo, otra característica que hace que los percentiles sean muy utilizados es la universalidad de su aplicación, ya que, a diferencia de los tipos de puntuaciones

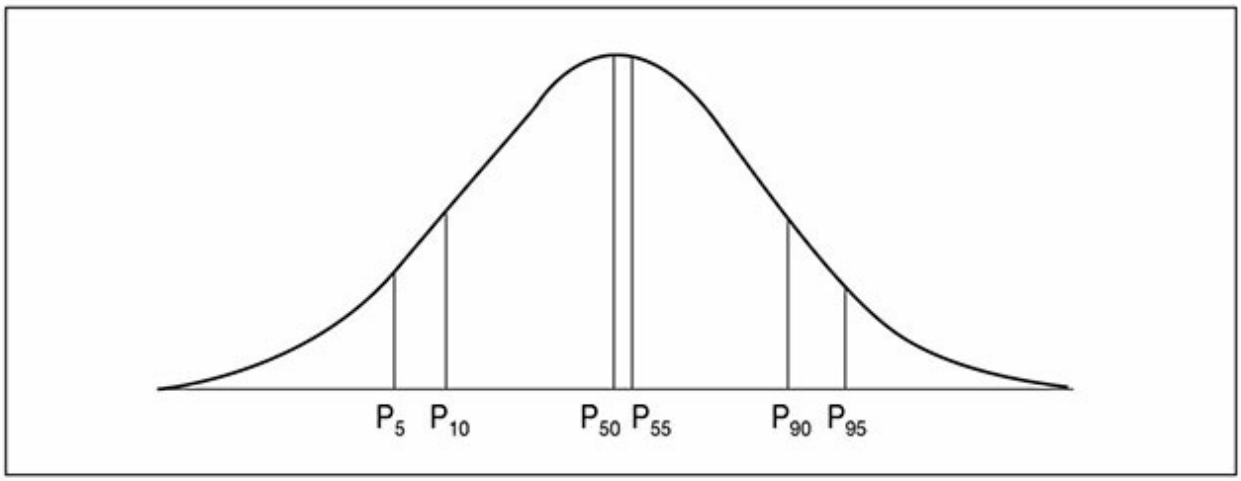
anteriormente revisados, pueden aplicarse a todas las edades y en todos los tipos de tests (de aptitudes, de personalidad y de rendimiento).

A pesar de estas ventajas, los percentiles presentan una limitación importante que los usuarios de tests deben tener siempre presente. En distribuciones que se aproximan a la curva normal —y esto ocurre con muchas variables educativas cuando el número de sujetos es relativamente alto—, la distribución no consta de partes iguales, con lo que la interpretación del percentil puede resultar algo engañosa. De esta forma, una diferencia muy grande en percentiles cerca de la mediana representa, con frecuencia, una diferencia pequeña en las puntuaciones directas, mientras que en los extremos ocurre justo lo contrario: pequeñas diferencias en percentiles equivalen a notables diferencias en puntuaciones directas. Esto sucede por el abultamiento que existe en la zona central de la curva (ver figura 4). Sabemos que ese abultamiento significa que una gran parte de los sujetos obtienen puntuaciones cercanas a la mediana -–que coincide con la media en la distribución normal—, mientras que el alisamiento de la curva en los extremos implica que sólo pocos sujetos obtienen puntuaciones extremas, tanto en el extremo inferior, como en el superior.

Esta idea se aprecia mejor con un ejemplo. Imaginemos las calificaciones (en la escala de 0 a 10) de una prueba objetiva (de elección múltiple) en un examen en la universidad, aplicado aproximadamente a 100 alumnos (para que los cálculos sean más sencillos). Es posible que la nota más baja sea un 0,5. Suponiendo que nadie más obtenga esta puntuación, el sujeto que ha obtenido esa nota será el percentil 1 de la distribución. La siguiente calificación podría ser un 1,5, ya que suelen ser escasas las puntuaciones tan bajas. Este segundo sujeto tendrá el percentil 2. Si comparamos el P_1 y el P_2, un solo punto de distancia entre los dos percentiles, equivale a 1 punto de diferencia en las puntuaciones directas de los dos alumnos del ejemplo.

FIGURA 4

DISTANCIA ENTRE PERCENTILES EN LOS EXTREMOS Y CENTRO DE LA CURVA



Es muy posible que, como suele suceder, en torno a la zona central de las puntuaciones (3,5 a 6,5) se sitúen muchos sujetos. Es decir, lo habitual es que un alto porcentaje de sujetos obtenga notas comprendidas entre ese intervalo central. Ello va a originar el mencionado abultamiento en la curva. De esta forma, entre el alumno que saca un 4,5 y el que saca un 5, es posible que haya muchos otros sujetos que han sacado puntuaciones comprendidas entre uno y otro valor (o sea entre 4,5 y 5). Por ello, y haciendo un cálculo aproximado, resulta posible que la puntuación directa 4,5 equivalga a un P_{40}, mientras que la puntuación 5, puede estar en el P_{50} de la distribución. A diferencia de lo que ocurría en el extremo inferior, entre la gran diferencia aparente de 10 unidades en percentiles, sólo existe una diferencia de 0,5 en las puntuaciones directas.

Por ello, los percentiles resultan una estimación gruesa de la posición relativa de la puntuación del sujeto en una distribución, o dicho de una forma más sencilla, nos permiten conocer el orden que ocupa el sujeto respecto al grupo normativo. Sin embargo, no nos permiten conocer con exactitud la *cuantía* de la diferencia entre su puntuación y la de otras personas. A pesar de ello, para una gran mayoría de procesos diagnósticos se utilizan las puntuaciones centiles ya que éstas ofrecen una información suficiente (PÉREZ JUSTE, 1991). Así, si no se pretende sino tener una información aproximada del orden de los sujetos, se puede preferir su uso por la facilidad con que son interpretadas y comprendidas por personas no expertas en el tema. Sin embargo, siempre debemos tener en cuenta que diferencias pequeñas en centiles —especialmente en el centro de la distribución— pueden no resultar relevantes.

6.2. Puntuaciones típicas o *z* y puntuaciones típicas derivadas

Las *puntuaciones típicas o puntuaciones z* presentan la ventaja de indicar la cuantía exacta de la diferencia que una puntuación mantiene respecto al promedio de un grupo. Para su cálculo es necesario conocer los valores de la media y desviación típica en el grupo de referencia, ya que hay que restar a la puntuación directa el promedio y dividir este valor por la desviación típica[23]. De esta forma, nos permiten saber cuántas unidades de desviación típica se diferencia cualquier puntuación respecto de la media.

Además de conocer las diferencias existentes entre diversas puntuaciones de un mismo test, las puntuaciones *z* son especialmente útiles para comparar resultados de los mismos sujetos en diferentes pruebas —para lo cual es necesario que las distribuciones tengan la misma forma, aunque, como hemos planteado, generalmente las puntuaciones de las variables con las que trabajamos en educación suelen tomar la forma de la curva normal—.

Supongamos que tenemos dos tests y los valores correspondientes a la media y desviación típica de las puntuaciones en ambos, tal y como se aprecia en el cuadro 6. Si quisiéramos saber quién obtiene una puntuación más alta en el test A, Ana o Carlos, es evidente que podemos responder a esta pregunta, simplemente observando sus puntuaciones directas. La puntuación de Ana es mayor que la de Carlos. Ahora bien, ¿cuánto mayor? Para poder estimar esto es necesario acudir a las puntuaciones *z*, de tal forma que podamos conocer la cuantía de la diferencia que ambas puntuaciones tienen respecto a la media en el test A. En el caso de Ana, su puntuación típica en el test A será de 1 (si aplicamos la fórmula para el cálculo de *z* antes mencionada), mientras que la de Carlos será de –1. Así, mientras Ana está una unidad de desviación típica por encima de la media, Carlos está a la misma distancia pero por debajo de la media.

En la misma situación presentada en el cuadro 6, podemos plantearnos comparar la puntuación de Ana y la de Carlos en ambos tests. Por ejemplo, ¿en cuál de los dos tests saca Ana una puntuación más alta? Si respondemos a esta pregunta basándonos en las puntuaciones directas, es posible que cometamos un error importante. En ese caso, aparentemente Ana obtiene mayor puntuación en el test A. Sin embargo, la respuesta correcta a esta pregunta es que Ana obtiene mayor puntuación en el test B. Si nos fijamos bien, en el test A, Ana está una unidad de desviación típica por encima de la media (recordemos que el valor de *z* era 1). Si calculamos la *z* de Ana en el test B, observamos que vale 1,5, lo que significa que su puntuación en el test B está una unidad y media de desviación típica por encima de la media siendo, por lo tanto, más alta. Con las puntuaciones de Carlos sucede lo mismo. Si su *z* en el test A era de –1, en el test B es de 0 (ya que su puntuación coincide, en este caso, con la media).

CUADRO 6

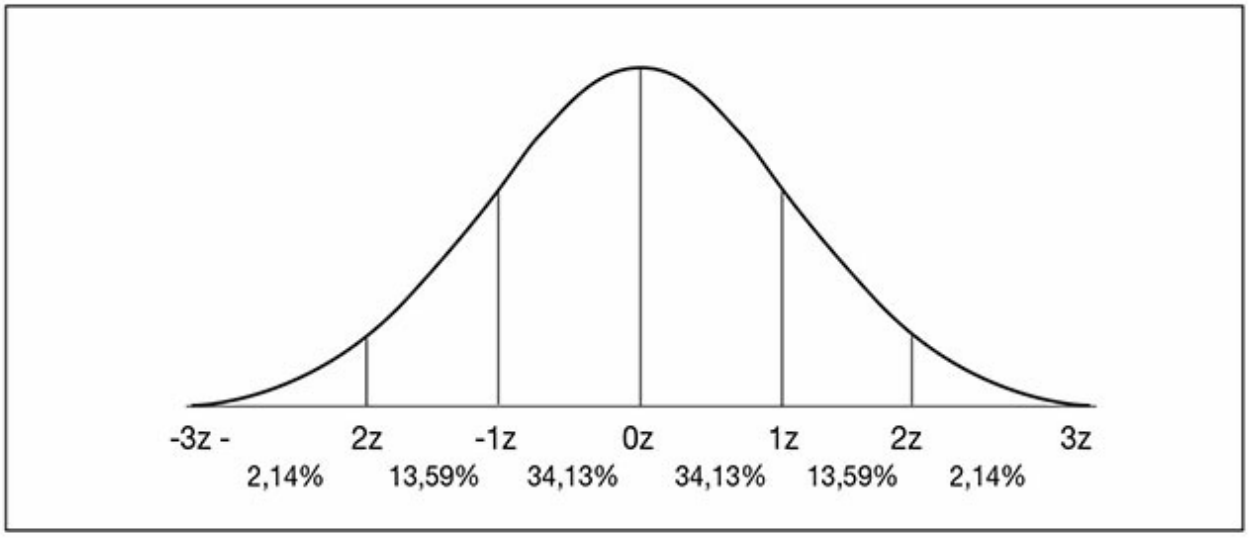
EJEMPLO DE APLICACIÓN DE LAS PUNTUACIONES *Z*

	TEST A	TEST B
Media	120	100
Desviación típica	15	10
	P.D. TEST A	**P.D. TEST B**
Ana	135	115
Carlos	105	100

Por tanto, la principal utilidad de las puntuaciones típicas es que nos permiten convertir las puntuaciones directas a una escala común de unidades iguales que pueden interpretarse en términos de la curva normal[24]. Esta escala guarda equivalencia con la de los percentiles, de tal forma que se sigue la relación mostrada en la figura 5. Mediante esta figura podemos observar que un 2% de los sujetos están dos desviaciones típicas por debajo de la media, aproximadamente un 16% (2,14% + 13,59%) están una desviación típica por debajo de la media, y así sucesivamente.

FIGURA 5

EQUIVALENCIAS ENTRE ALGUNAS PUNTUACIONES *Z* ENTERAS Y LOS PERCENTILES



Por lo que llevamos visto, se deduce que las puntuaciones *z* presentan múltiples ventajas. Podemos preguntarnos si tienen algún inconveniente, a lo que tendríamos que

responder que sí. Más que un inconveniente propiamente dicho, se trata de una «incomodidad» o dificultad a la hora de ser interpretadas. Como se aprecia en la figura 5, las puntuaciones *z* suelen oscilar entre –3 y +3 (llegando rara vez a igualar o superar, en valores absolutos, estos límites). Por ello, es habitual que tomen valores decimales y negativos, lo que plantea dificultades para conseguir una interpretación del valor de *z* que resulte clara e inteligible. Si le decimos a un sujeto que está 1,15 unidades de desviación típica por debajo de la media, ¿qué le queremos decir?

Para solventar esta dificultad existen diferentes tipos de *puntuaciones derivadas* basadas en el cálculo de *z* que nos permiten evitar las puntuaciones negativas y los decimales. Estas puntuaciones se obtienen haciendo una transformación lineal de las puntuaciones típicas, lo que permite que la nueva distribución tenga las mismas propiedades que las *z*. Generalmente, la transformación responde al tipo expresado en la siguiente fórmula: Puntuación típica derivada = Constante A + Constante B (*z*). El nuevo conjunto de puntuaciones obtenido tras esta transformación, tendrá como media el valor de la constante A y como desviación típica el de la constante B.

Son muchas las puntuaciones típicas derivadas que pueden obtenerse con este procedimiento, ya que mediante él se puede obtener un conjunto de puntuaciones «a gusto del consumidor». Todo dependerá del margen con que queramos que oscile el nuevo conjunto de puntuaciones (de 1 a 10, de 1 a 100, etc.). Dado que son muchos los tipos de puntuaciones típicas derivadas, aquí sólo vamos a ver dos de ellas que son las más usadas: las estaninas o eneatipos y las puntuaciones T.

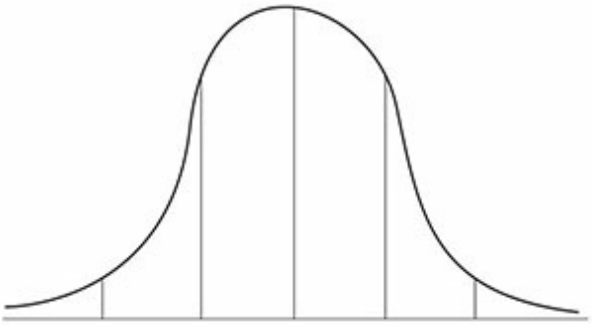
Las *estaninas o eneatipos* se obtienen de multiplicar la puntuación *z* por la constante 2 y sumarle la constante 5 ($E=5+2z$). De esta forma, se obtiene un conjunto de puntuaciones que oscilan entre 1 y 9 y que tiene su media en 5 y su desviación típica en 2, conservando todas las ventajas de las puntuaciones típicas. Los valores decimales se desestiman (ya hemos indicado que la finalidad de usar puntuaciones derivadas es evitar decimales y números negativos), de forma que podemos interpretar fácilmente la puntuación en eneatipo de un sujeto según se aproxime más o menos al extremo más bajo (1), a la media (5) y al extremo más alto (9).

Muy similar es el procedimiento que se aplica para obtener las *puntuaciones T*. En este caso las constantes son 50 y 10 ($T=50+10z$), de forma que obtenemos un conjunto de puntuaciones con media en 50 y con una desviación típica de 10, que pueden oscilar aproximadamente entre 20 y 80, sin tomar valores decimales o negativos que dificulten la interpretación.

En la figura 6 sintetizamos las relaciones que existen entre los diferentes tipos de puntuaciones transformadas que hemos visto en este apartado. El establecimiento de estas relaciones es posible siempre y cuando las mediciones realizadas adopten la forma de la curva normal, que es aquí el marco desde el cual se establece la comparación.

FIGURA 6

RELACIÓN ENTRE DIFERENTES TIPOS DE PUNTUACIONES



Desviaciones típicas	-3	-2	-1		0	1		2	3
Puntuaciones z	-3	-2	-1		0	1		2	3
Percentiles	0,1	2	16		50	84		98	99,9
Puntuaciones T	20	30	40		50	60		70	80
Estaninas	1	2	3	4	5	6	7	8	9

(Basado en PÉREZ JUSTE, 1991: 317)

7. GARANTÍAS TÉCNICAS Y CIENTÍFICAS EN LOS INSTRUMENTOS DE MEDIDA

Al ser los instrumentos que más se aproximan a la medición, es en los tests y pruebas objetivas donde adquieren mayor importancia y sentido las características técnicas de validez y fiabilidad, que han sido abordadas brevemente en otros capítulos de este libro. Por ello, hemos optado por profundizar aquí en estos conceptos, así como en los procedimientos que se usan para su estimación.

En este sentido, conviene aclarar que consideramos que las cuestiones referidas a la formulación y tratamiento estadístico de los procedimientos de estimación de la validez y fiabilidad no entran dentro del alcance de este libro. En su lugar, preferimos centrarnos en los dos aspectos que sí consideramos clave para la formación de los profesionales del diagnóstico: por un lado, los conceptos mismos de validez y fiabilidad y, por otro, los procedimientos que, a nivel general, nos permiten su estimación.

No obstante, antes de adentrarnos en la conceptualización de la validez y la fiabilidad, hemos creído necesario hacer referencia a dos cuestiones que están relacionadas con las características técnicas de los instrumentos de medida. Nos referimos, por un lado, a un conjunto de tendencias en la respuesta de los sujetos que afectan a la calidad de la medida y, por otro, al análisis de ítems, como conjunto de procedimientos que contribuyen a la mejora de los instrumentos.

A continuación, empezamos abordando estos dos temas para proseguir posteriormente con la cuestión de la validez y la fiabilidad de los instrumentos de medida.

7.1. Tendencias de respuesta que afectan a la medida

En este apartado nos vamos a referir a un conjunto de tendencias que los sujetos pueden manifestar a la hora de responder a los tests no cognoscitivos y escalas de actitudes. En principio, los estilos de respuesta que vamos a describir no afectan a otro tipo de pruebas como son las cognoscitivas (tanto las destinadas a medir aptitudes como las destinadas a medir rendimiento). En estos tipos de pruebas, las respuestas de los sujetos pueden calificarse como de «correctas» o «incorrectas», cosa que no sucede en las pruebas afectivas, donde las respuestas que el sujeto da son opiniones o manifestaciones que éstos hacen de sus valores, actitudes, rasgos de personalidad, etc. No obstante, en algunos tipos de pruebas de rendimiento también pueden incidir estas tendencias de respuesta. De hecho, los primeros estudios sobre la aquiescencia fueron realizados por Cronbach[25] en pruebas de verdadero-falso, demostrando este autor que muchos sujetos tienden a elegir la opción «verdadero» cuando dudan sobre la respuesta correcta.

Nos centraremos en dos tendencias de respuesta que afectan especialmente a las pruebas de personalidad y a las escalas de actitudes, frente a otros instrumentos de evaluación. Estas dos tendencias son la aquiescencia y la deseabilidad social. Sin embargo, no son las únicas que pueden incidir en este tipo de pruebas. Existen otras en las que no vamos a profundizar, dado que han sido mencionadas ya en el capítulo correspondiente a la observación. Allí hablábamos de la tendencia central, de la tendencia a dar puntuaciones muy positivas o muy negativas, y del efecto de halo.

Se denomina *aquiescencia* a la tendencia de algunos sujetos a responder afirmativamente con independencia del contenido, mostrando incluso acuerdo con afirmaciones de signos opuestos. En realidad, se trata de un modo de responder en las pruebas no cognoscitivas que invalida la información que aparentemente está presente en la respuesta, ya que una respuesta afirmativa no indicaría necesariamente acuerdo con la afirmación contenida en el reactivo, sino que sería simplemente una manifestación de dicha tendencia de respuesta.

Por lo general, para paliar sus efectos se suele recomendar que las pruebas que se utilizan tengan ítems positivos (favorables hacia la actitud, rasgo o característica que se quiere medir) y negativos (desfavorables hacia la misma). Hay que aclarar, no obstante, que los ítems negativos se formulan preferentemente en términos positivos (y no necesariamente utilizando adverbios como «no» o «nunca»). Ya hemos hecho referencia a una cuestión parecida cuando abordábamos la formulación de ítems en las pruebas objetivas.

MORALES (2000) hace un análisis bastante exhaustivo del problema de la aquiescencia. Entre las principales ideas que este autor destaca en torno al tema, podemos resaltar las siguientes:

— Hay suficiente evidencia empírica como para afirmar que, en ocasiones, se manifiesta de manera más o menos consistente la tendencia de algunos sujetos a responder afirmativamente (o a estar de acuerdo) independientemente del contenido de los ítems, e incluso con aparente contradicción.

— La aquiescencia que puede manifestarse en una escala de actitudes es distinta a la que puede estar presente en las respuestas a un test de personalidad. Diversos tipos de instrumentos estimulan la aquiescencia en sujetos distintos.

— La ambigüedad de los ítems y la dificultad en entenderlos es posiblemente el estímulo más claro que incita a algunos sujetos a dar respuestas aquiescentes. Lo mismo puede decirse de la relevancia percibida del instrumento. La aquiescencia tiene más que ver con la calidad del instrumento y su adecuación a los sujetos a quienes va dirigido, que con características personales de los que responden.

— Por lo general, las escalas que incluyen aproximadamente idéntico número de ítems positivos y negativos se prestan menos a estimular la aquiescencia y se responden con más cuidado. Aunque no la evitan, permiten comprobarla

examinando las correlaciones entre ítems de signo opuesto; este tipo de escalas exigen una definición más matizada del constructo y una formulación más cuidadosa de los ítems. En la construcción de instrumentos, como mínimo, debe evitarse que todos los ítems estén en una misma dirección.

Aunque existen diferentes métodos que nos permiten verificar si se da o no aquiescencia en una situación concreta (escalas específicamente diseñadas para medir esta tendencia), lo mejor es prevenirla elaborando instrumentos en los que la incidencia de este estilo de respuesta sea mínima. Para autores como KERLINGER (1975), el influjo potencial de la aquiescencia es irrelevante si se utilizan instrumentos razonablemente bien hechos.

Otro importante problema que afecta a la medición en las pruebas no cognoscitivas es el de la *deseabilidad social*, el cual nos remite a la pregunta siguiente: ¿hasta qué punto el sujeto es sincero a la hora de contestar? o ¿sus respuestas están distorsionadas por lo que considera socialmente deseable? Las respuestas del sujeto pueden estar influidas por su deseo de dar una buena imagen de sí mismo, lo cual se refleja en su tendencia a responder de una manera socialmente adecuada. Esta tendencia, no obstante, es difícil de conceptualizar porque en muchas ocasiones, el propio sujeto no es consciente de su falta de sinceridad, sino que responde desde su yo ideal, es decir, desde lo que desea o aspira a ser. Así, un sujeto que se autodescribe (bien sus rasgos de personalidad, bien sus actitudes ante un tema) en términos preferentemente positivos, no necesariamente está mintiendo: la imagen de sí mismo que presenta puede responder a la realidad o a la imagen de sí mismo que tiene.

Al igual que ocurre con la aquiescencia, existen diferentes instrumentos que permiten conocer si las respuestas están influidas por la deseabilidad social de los ítems más que por el contenido de los mismos. No obstante, nuevamente, lo importante es planificar los instrumentos de manera que se minimice la incidencia de esta tendencia de respuesta. MORALES (2000) informa de dos enfoques utilizados para controlar la deseabilidad social: por un lado, cambiando el sistema de medición por uno más resistente al engaño, y por otro, mejorando el sistema tradicional con determinadas precauciones que faciliten la respuesta sincera.

En el primer enfoque, se trata de sustituir los tradicionales ítems unitarios —en los que únicamente subyace un rasgo y se responden de manera independiente, uno a uno— por ítems de elección forzada. En un apartado anterior de este capítulo presentamos un ítem del SIV que puede servir ahora como ejemplo de ítem de elección forzada; en este caso, el sujeto se ve forzado a escoger entre valores con igual nivel de deseabilidad social. No existen respuestas más «deseables socialmente» que otras.

Además de la elección forzada, se puede controlar la deseabilidad social utilizando las pruebas convencionales, pero poniendo en práctica lo que MORALES (2000) denomina como *medidas preventivas* que van a facilitar la sinceridad de la respuesta. Según este autor, estas medidas preventivas son suficientes en la mayoría de las situaciones. En

primer lugar, es necesario formular los ítems de manera que no susciten fácilmente respuestas socialmente deseables, intentando que sean neutros en deseabilidad social; otra recomendación se refiere a conservar el anonimato, lo cual favorece la sinceridad; una tercera sugerencia se refiere a la necesidad de establecer un clima de cercanía y confianza con los que van a responder, pidiendo abiertamente que las respuestas sean sinceras.

Aunque la aquiescencia y la deseabilidad social son problemas serios que afectan a la medición de factores no cognoscitivos y merecen, por ello, un tratamiento más exhaustivo del que se les ha dado aquí, esperamos haber dejado patente, al menos, la importancia de cuidar la redacción de los ítems en la elaboración de todo instrumento de medida.

7.2. Análisis de ítems

Si bien el apartado anterior estaba referido a las pruebas afectivas, este apartado se centra en un conjunto de operaciones que facilitan la mejora, fundamentalmente, de las pruebas cognoscitivas, aunque algunos de estos análisis pueden aplicarse, en ocasiones, a otro tipo de pruebas.

En términos generales, el análisis de ítems es el proceso consistente en examinar las respuestas dadas por los sujetos a cada ítem de un test o prueba, con el propósito de juzgar la calidad de los reactivos. En la construcción de una prueba, el objetivo general es llegar a un test de mínima longitud que produzca puntuaciones con el mayor grado de fiabilidad y validez. Esto se consigue a través del análisis de sus ítems. El proceso consiste en estimar las propiedades estadísticas de los ítems que más van a influir en la fiabilidad y validez del test resultante.

Entre sus aplicaciones no sólo está el elegir los mejores ítems para ser incluidos en la prueba, sino que, en el caso de las pruebas objetivas de conocimiento, el análisis de ítems nos va a permitir conocer empíricamente las propiedades de los reactivos, a fin de reutilizarlos en ocasiones sucesivas, seleccionando los ítems, por ejemplo, según su nivel de dificultad.

Las características o propiedades de los ítems que se suelen analizar se calculan a través de determinadas formulaciones estadísticas que permiten conocer aspectos como:

— El *índice de dificultad* o proporción de sujetos que han contestado correctamente al ítem.

— El *índice de discriminación* o grado en el que el ítem permite detectar las diferencias individuales en la ejecución del test.

— El *índice de homogeneidad o* consistencia interna que indica si un sujeto que contesta de una determinada manera en el test, contesta de la misma forma en un

ítem particular.

— El *índice de validez* o correlación que mantiene cada ítem con un criterio externo.

Además de índices como los mencionados, el análisis de los distractores aporta una importante información sobre los reactivos. En aquellas pruebas en las que los ítems responden al formato de elección múltiple, nos interesa conocer la eficacia de las alternativas incorrectas que se presentan junto con la opción verdadera. Desde el punto de vista teórico, lo deseable es que todos los distractores atraigan por igual al sujeto que no conoce la respuesta. Para el análisis de distractores, por lo general, se forman dos grupos de sujetos: el grupo de nivel superior en el que se agrupan los sujetos con puntuaciones más altas, y el grupo de nivel inferior, del que forman parte los sujetos con puntuaciones más bajas. Se contabilizan las frecuencias o porcentajes de elección de cada alternativa en cada uno de los dos grupos, considerando que (CABRERA y ESPÍN, 1986):

— Un buen distractor debe ser más elegido por los sujetos del grupo inferior que por los del grupo superior.

— Un distractor más elegido por los del grupo superior es deficiente y necesita ser mejorado.

— Un distractor apenas elegido —ni por uno u otro grupo— es ineficaz.

Otra importante línea de investigación en torno a los ítems se centra en analizar si éstos funcionan de tal forma que tiendan a favorecer a un grupo (grupo de referencia) sobre otro (grupo focal). En ese contexto se dice que el ítem o ítems presenta/n un funcionamiento diferencial (FDI) en cada uno de los grupos comparados. Este tema está suscitando numerosos trabajos de investigación, por las importantes implicaciones que tiene el hecho de que los tests puedan presentar sesgos en la evaluación. En GIL, GARCÍA y RODRÍGUEZ (1998) se puede encontrar un estudio comparativo de algunos de los métodos usados para la detección del funcionamiento diferencial de los ítems.

7.3. Características técnicas de los instrumentos de medida

Cualquier instrumento de evaluación y medida debe reunir dos características: la validez y la fiabilidad. Dicho de otra forma: un instrumento de medida debe medir aquello que se propone medir (*validez*) y además medirlo con precisión (*fiabilidad*).

Si entendemos por precisión el grado de correspondencia entre la medida empírica (u observada) y la magnitud real (la medida verdadera), la fiabilidad, en cuanto precisión, es un concepto que no puede ser conocido empíricamente en términos absolutos (*fiabilidad absoluta*). Es obvio que cualquier información de que disponemos de la puntuación verdadera proviene precisamente de la medida que hemos realizado y, por tanto, desconocemos la verdadera magnitud de lo que hemos medido. Por esta razón, se

recurre a otro concepto de fiabilidad (*fiabilidad relativa*), el cual nos permite conocer de forma indirecta si estamos midiendo con cierta precisión.

La fiabilidad relativa hace referencia a la estabilidad o consistencia de las medidas. Aquí el término «fiabilidad» corresponde a su auténtico significado; es fiable aquello que se mantiene constante en situaciones semejantes (y, por tanto, nos podemos fiar del valor obtenido), o aquello que es consistente (en tanto que constante en todas sus partes o dimensiones). Un test es fiable si cada vez que se aplica a los mismos sujetos proporciona los mismos resultados. Y es a partir de la constancia como inferimos la precisión del instrumento, que es el concepto que realmente nos interesa; cuanto más constante, más preciso. Esto es así porque suponemos que la puntuación verdadera estará dentro del rango de la variabilidad observada, y deduciremos que estamos midiendo con mayor precisión, cuanto menor sea la fluctuación entre las puntuaciones observadas en igualdad de circunstancias.

En toda medición incide un componente de error que es aleatorio. Los errores aleatorios o errores de medición son consecuencia de los múltiples factores que inciden en la medida y que introducen en la misma una cierta variabilidad. Variables como el estado de ánimo, cansancio, grado de adivinación, errores de codificación, etc. pueden afectar a los resultados de la prueba aplicada. Tales errores hacen referencia a la precisión de las medidas y entran dentro del terreno de la fiabilidad. El componente de error que hay en la medida posee las características siguientes:

1. El error medio, en caso de sucesivas aplicaciones del instrumento, es cero. Es decir, los errores actúan unas veces en un sentido y otras en el contrario, de forma que quedan compensados entre ellos y, a la larga, su suma valdrá cero.
2. El error no está correlacionado, lo que indica que el error correspondiente a una aplicación del test no correlaciona con el error correspondiente a otra aplicación.
3. El error tampoco correlaciona con la puntuación verdadera.

Bajo estos supuestos, por tanto, la puntuación verdadera de un sujeto en un test puede definirse teóricamente como la puntuación media resultante de efectuar mediciones repetidas del mismo test al mismo individuo.

El concepto de *validez*, aparentemente simple si atendemos a la definición ofrecida anteriormente («una medida es válida si mide aquello que pretendemos medir»), es algo más complejo. Por ejemplo, si estuviésemos midiendo la altura de una persona, podríamos suponer que la cinta métrica es válida para medir longitudes. Sin embargo, la longitud es un fenómeno directamente observable tal cual se da en la naturaleza. En educación y en psicología, frecuentemente trabajamos con constructos que no se pueden observar de forma directa. Así, cuando construimos una prueba que pretende medir la inteligencia, ¿cómo podemos estar seguros de que medimos este rasgo y no otro?, ¿qué garantías tenemos de que no estamos midiendo otros rasgos (fluidez verbal, razonamiento...)?

A diferencia de la fiabilidad, la validez obedece más a consideraciones teóricas que a cuestiones empíricas, ya que hay una dimensión cualitativa de por medio; se trata de saber *qué* medimos, mientras que con la fiabilidad nos preocupamos por el grado de precisión de la medida, el cual se puede cuantificar en la constancia de los resultados.

La validez, al igual que la fiabilidad, también está amenazada por una fuente de error. Se trata del error que denominamos conceptual. Los errores conceptuales son consecuencia, no tanto de medir mal una determinada variable, como de medir equivocadamente otra en su lugar. Este tipo de errores, exige que los planteamientos teóricos de partida de un instrumento de medida tengan que ser muy rigurosos en torno a la naturaleza de las variables.

De cualquier forma, aunque los conceptos de validez y fiabilidad son diferentes, existe una cierta relación entre ellos. Un buen nivel de fiabilidad no implica necesariamente una buena validez, pero una mala fiabilidad sí implica una deficiente validez. Por el contrario, una buena validez, exige necesariamente una fiabilidad previa igualmente buena.

Una vez introducidos de forma general los conceptos de validez y fiabilidad, pasamos a continuación a profundizar en los procedimientos que se usan para asegurar ambas características técnicas.

7.3.1. Fiabilidad

Como ya hemos comentado, la fiabilidad se conoce de forma indirecta, entendiendo que se muestra en la constancia o estabilidad de las medidas aportadas por un instrumento en situaciones semejantes. Así, un test será fiable si tras diferentes pasadas a los mismos individuos ofrece resultados parecidos. En términos estadísticos, la fiabilidad de un test quedará reflejada mediante la correlación de las puntuaciones empíricas obtenidas en dos situaciones semejantes. Cuanto mayor sea la correlación, más parecidas o constantes serán tales puntuaciones y, en consecuencia, mayor será su *coeficiente de fiabilidad*.

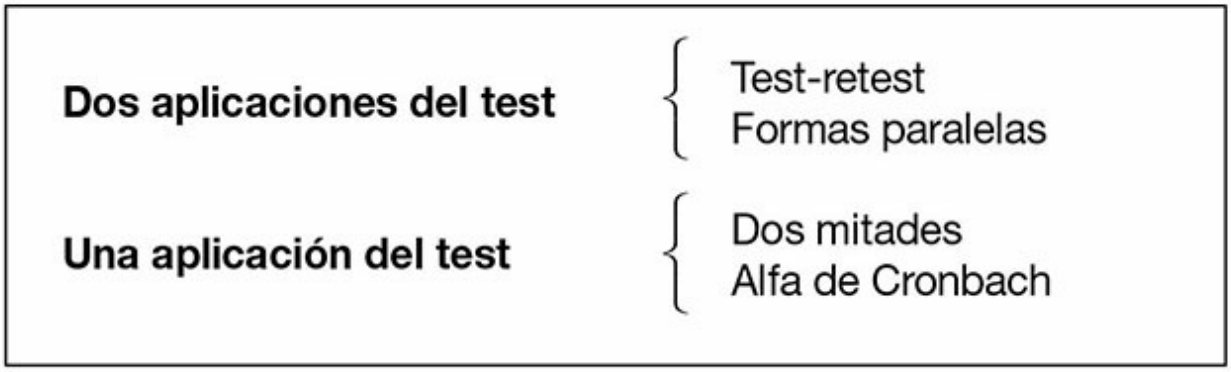
Esto es más complejo de lo que parece. En el ámbito educativo y psicológico, los atributos que pretendemos medir están sujetos a cambios. En este sentido, no debemos confundir la constancia de las mediciones con la constancia de las puntuaciones (lo que medimos). Una prueba será más fiable si es capaz de reflejar tal cambio, sin confundirlo con la falta de precisión de la medida.

Pese a estas dificultades, dentro de la *Teoría Clásica de los Tests* (TCT) se han planteado diferentes procedimientos de cálculo del coeficiente de fiabilidad, todos ellos basados en el coeficiente de correlación de Pearson. En la figura 7 se presenta un resumen de los principales procedimientos utilizados para el cálculo del coeficiente de fiabilidad. Como puede observarse, se distingue un primer grupo de estrategias que se basan en *dos aplicaciones del test*, bien sea el mismo test (método test-retest), o dos tests equivalentes (método de formas paralelas). El segundo grupo de procedimientos se basa en una *única aplicación del test*, dividiendo éste en dos mitades (método de las dos

mitades) o considerando cada uno de los ítems de que se compone como elementos independientes (alpha de Cronbach).

FIGURA 7

COEFICIENTES DE FIABILIDAD



(CAMACHO y SÁNCHEZ, 1997:46)

Mediante los procedimientos de dos aplicaciones del test se obtiene un coeficiente de fiabilidad como *estabilidad* de las medidas, mientras que con una única aplicación, lo que medimos es la *consistencia interna* de las medidas. En principio, ninguno de estos procedimientos es más adecuado que otro. Dependiendo de las circunstancias, será aconsejable el uso de uno o varios de ellos. A continuación, haremos un breve repaso de todos los procedimientos mencionados en la figura 7, describiendo sus características principales sin entrar en la formulación estadística que permite su cálculo.

1) El método del test-retest

Consiste básicamente en aplicar el mismo test al mismo grupo de sujetos en dos ocasiones distintas, calculando el coeficiente de correlación de Pearson entre las dos series de puntuaciones. Una alta correlación —próxima a 1— implica que el test presenta estabilidad al medir el rasgo y, por lo tanto, una alta fiabilidad.

Este método es adecuado cuando se miden rasgos que varían poco en el tiempo y procurando que el intervalo temporal entre ambas aplicaciones no sea excesivamente corto (podría existir recuerdo), ni excesivamente largo (en cuyo caso, podrían afectar los procesos madurativos). El coeficiente de fiabilidad que se obtiene se denomina *coeficiente de estabilidad*, en tanto que refleja el grado en que las medidas se mantienen estables a lo largo de las dos aplicaciones del test.

2) El método de las formas paralelas

Este método consiste en la aplicación al mismo grupo de sujetos de dos versiones distintas (dos formas paralelas) de la misma prueba. Las dos versiones de la prueba, para

poder ser consideradas como verdaderamente «paralelas», han de medir los mismos aspectos con el mismo tipo de cuestiones (aunque éstas, por supuesto, no pueden ser iguales). En general, obtener dos formas exactamente paralelas es una cuestión difícil, por lo que es más razonable hablar de dos formas alternativas, entendiendo que ambos tests no son por completo equivalentes, sino dos «intentos» de que así lo sean.

En este procedimiento, la memoria o el recuerdo no afecta a la respuesta de los sujetos, por lo que sería el procedimiento idóneo para la estimación de la fiabilidad, si no fuera por la dificultad de elaborar dos pruebas realmente equivalentes. En caso de que las pruebas no lo sean totalmente, resulta difícil distinguir en la puntuación verdadera el cambio de la falta de fiabilidad.

3) El método de las dos mitades

Este procedimiento se basa en la aplicación del mismo test a un mismo grupo de sujetos, dividiendo posteriormente el test en dos mitades, es decir, la mitad de los ítems configura uno de los tests y la otra mitad el otro. Si las puntuaciones de ambas mitades presentan un alto coeficiente de correlación de Pearson, el test tendrá un buen coeficiente de fiabilidad.

Para constituir dos partes lo más equivalentes posible, el procedimiento más usado es seleccionar los ítems pares para la primera mitad y los impares para la segunda. También se pueden elegir aleatoriamente los ítems que compondrán cada mitad. No resulta conveniente utilizar un procedimiento que divida al test en dos mitades que puedan no ser equivalentes, como sucedería si tomáramos la primera mitad de los ítems para configurar uno de los tests, y la segunda mitad para el otro. En este caso, los ítems podrían estar ordenados por dificultad, y además la fatiga incidiría más en la segunda parte que en la primera, y como consecuencia de ello, las dos mitades no serían equivalentes.

Este procedimiento presenta alguna de las ventajas de los métodos anteriores, al tiempo que supone un ahorro de tiempo y esfuerzo y, más importante aún, al no haber dos aplicaciones eliminamos el efecto del intervalo temporal y de memoria o recuerdo. El principal inconveniente que presenta proviene del hecho de que el valor del coeficiente de fiabilidad obtenido depende de cómo han quedado repartidos los ítems en cada una de las mitades. Posiblemente, la correlación entre los ítems pares e impares será diferente que la correlación entre las dos mitades si éstas se seleccionan por un procedimiento aleatorio. La única forma de garantizar que los coeficientes de fiabilidad posibles coincidan consiste en utilizar diferentes criterios para la división en las dos mitades y, al mismo tiempo, exigir paralelismo en ambas.

Los procedimientos de Spearman-Brown, Rulon y Flanagan y Guttman son aplicaciones estadísticas que permiten el cálculo del coeficiente de correlación de Pearson entre ambas mitades cuando se aplica el agrupamiento de los ítems pares e impares.

4) El coeficiente Alpha de Cronbach

Como acabamos de sugerir, el método de las dos mitades presenta algunos inconvenientes para determinar la fiabilidad de un test en tanto que consistencia interna. Entre los procedimientos que intentan aproximarse a la consistencia interna de una prueba con un único coeficiente, el más conocido y utilizado es el coeficiente alpha de Cronbach. Este procedimiento se basa en la correlación media entre todos los ítems de un test. Expresa el grado en que todos los ítems miden el mismo rasgo y lo miden con precisión.

Un valor alto de este coeficiente indica que la correlación promedio entre los ítems es grande. Sin embargo, este índice depende de otros factores tales como el número de ítems en cuestión o el número de factores subyacentes en cada ítem. Por esto hay que completar la información que nos ofrece el coeficiente alpha con otros procedimientos estadísticos. Por ejemplo, si en una prueba existen varios sub-tests, es conveniente calcular el coeficiente de fiabilidad para cada una de las partes, para finalmente calcular la fiabilidad global del test como una combinación lineal de tales sub-tests.

7.3.2. Validez

Como ya se ha indicado, la validez garantiza que medimos aquello que nos proponemos medir, y no otra cosa parecida o relacionada. La validez de un instrumento es la utilidad de sus puntuaciones y la adecuación de ellas a las decisiones que de las mismas deriven (AERA/APA/NCME, 1985).

A lo largo de los años, la definición de validez ha pasado de poner el énfasis en el test en sí mismo a centrarse cada vez más en la interpretación que se deriva de sus puntuaciones. En sentido estricto, no se valida un test, ni siquiera las puntuaciones obtenidas en el mismo, sino las consecuencias que se derivan de la interpretación de dichas puntuaciones o el uso que se hace de las mismas (MESSICK, 1995) o, en palabras de CRONBACH y MEEHL (1955:297), «no se valida un test, sino sólo un principio desde el cual hacer inferencias».

Este nuevo concepto de validez plantea algunos cambios respecto a su formulación y estimación. La validez es un proceso —no un producto— de acumulación de evidencias que justifiquen las inferencias derivadas (GARCÍA, GIL y RODRÍGUEZ, 1996). No hay un único coeficiente, ni tampoco un único método para «demostrar» la existencia de validez. Por el contrario, la validación de un instrumento consiste en el *proceso que permite ir depurando, delimitando y perfeccionando cada vez más el instrumento* que usamos para una finalidad determinada (CAMACHO y SÁNCHEZ, 1997). Puede por ello decirse que la validez no es una propiedad general de los instrumentos de medida, sino su utilidad para un objetivo concreto. La validez, por tanto, debe entenderse como un proceso continuo que requiere la acumulación de datos procedentes de distintas investigaciones empíricas y procesos lógicos (CABRERA y ESPÍN, 1986).

Otra cuestión de gran importancia en relación con el concepto de validez y los cambios que dicho concepto ha experimentado a lo largo del tiempo, se refiere al número y tipos de validez definidos. Así, de los tres tipos de validez comúnmente aceptados (AERA/APA/NCME, 1985), de contenido, referida a un criterio y de constructo, es precisamente esta última la que mayor interés suscita, siendo habitual la postura que defiende que engloba a los demás tipos de validez (LOEVINGER, 1957; MESSICK, 1980).

De esta forma, si la validez de constructo es el grado en que un instrumento mide el constructo o rasgo que pretende medir, cualquier indicio de validez (ya sea referida a un criterio externo o de contenido) puede ser considerado como evidencia de la validez del constructo. Ésta implica no sólo obtener evidencias de que los ítems son una medida suficientemente representativa del universo de ítems que sirven para medir el dominio en cuestión (validez de contenido), sino que además implica el estudio de la relación del constructo a medir con otros constructos o rasgos relacionados con el mismo y que forman parte de una misma teoría (validez referida a un criterio). Esto es lo que se ha denominado como *perspectiva unificadora de la validez* frente a la visión trinitaria de la misma (ELLIS y BLUSTEIN, 1991).

Esta perspectiva unificadora nos lleva a una importante conclusión: la validez de contenido, de criterio y de constructo no son diferentes tipos de validez, sino diferentes facetas o aspectos de la misma. Estas tres facetas no son sino caminos distintos para obtener las evidencias de la validez; se trata de aproximaciones a la misma, siendo único el concepto de validez (GARCÍA, GIL y RODRÍGUEZ, 1996).

Teniendo en cuenta estas importantes consideraciones, a continuación vamos a referirnos a las diferentes facetas de la validez, entendidas como vías de aproximación a la misma que contribuyen a reunir evidencias en función de los objetivos planteados con la medida.

1) Validez de contenido

La validez de contenido se refiere al grado en que la muestra de ítems que componen el test es representativa del universo o dominio al que el test está referido. Se basa en la evidencia representacional de los reactivos o ítems con respecto al dominio educativo, psicológico o de conducta que se pretende medir (MARTÍNEZ ARIAS, 1995). Requiere, por tanto, una gran rigurosidad para derivar de objetivos y contenidos el mapa ponderado de indicadores tipo que ha de constituir la prueba. Este aspecto de la validez hace referencia a la *relevancia* del contenido del instrumento y a la *representatividad* de sus ítems respecto al dominio. La validez de contenido es, por tanto, una validez fundamentalmente teórica o de representación.

Para evaluar la validez de contenido de un test o prueba, CROCKER y ALGINA (1986) proponen un proceso de elaboración de los contenidos que pasa por la revisión de expertos:

1. Definir el dominio o universo de contenidos objetivo del test.
2. Seleccionar un conjunto de expertos cualificados en dicho dominio.
3. Proporcionar un marco estructurado para el proceso de emparejar los ítems con el dominio referido.
4. Recogida y resumen de los datos resultantes del proceso de emparejamiento mencionado.

A partir de la existencia previa de una tabla de especificaciones —en los términos descritos en apartados anteriores—, se puede establecer la representatividad de los ítems. Además, si se proporciona a los expertos la lista de los objetivos y contenidos relativos al rasgo a evaluar, y los ítems elaborados para cada objetivo, éstos pueden juzgar y cuantificar la relevancia de cada ítem (MARTÍNEZ ARIAS, 1995).

Para cuantificar la relevancia de los ítems, HAMBLETON (1980) propone utilizar una escala de 1 a 5, donde 1 indica un mal ajuste del ítem al objetivo definido y 5 un ajuste perfecto entre ítem y objetivo. De esta forma, la relevancia de cada ítem será definida por la media o mediana de las puntuaciones dadas por los expertos.

2) Validez referida a un criterio

La validez relativa a un criterio consiste en comprobar el grado de asociación que existe entre el rasgo o constructo que mide el test con otro/s constructo/s que se supone/n relacionado/s. En esencia, consiste en comparar la medida que ofrece el test con un estándar del que se tiene constancia previa de su validez. A diferencia de la validez de contenido, la validez de criterio pone el énfasis en conocer en qué medida la prueba sirve para predecir una conducta, o bien como medida sustitutiva de otra ya existente.

El rasgo a validar y el criterio pueden ser medidos a la misma vez (*validez concurrente*) o en momentos distintos, siendo posterior la medida del criterio (*validez predictiva*). Un problema esencial que afecta a la validez de criterio es la selección de los mismos criterios, cuando la lógica de la validez implica necesariamente que el criterio sea válido. En ocasiones, no se dispone de los criterios idóneos, hecho que no siempre se cuestiona a pesar de su importancia. La atención de los constructores de pruebas se centra más en el nuevo instrumento que en verificar la calidad de los criterios escogidos (MORALES, 2000).

El valor de la validez de criterio de una prueba se expresa mediante el coeficiente de correlación de Pearson (que se denomina en este contexto *coeficiente empírico de validez*) entre la variable objeto de estudio y la variable considerada como criterio. Obsérvese que tan sólo nos importa que exista una fuerte correlación entre el test y el criterio de referencia. El test será *válido* en tanto se relacione con un criterio externo y sea eficaz en su predicción.

3) Validez de constructo

Anteriormente hemos hecho referencia a la importancia que tiene la validez de constructo. Según CLARK y WATSON (1995), las medidas más precisas y eficientes son aquellas en las que se establece la validez de constructo, en tanto que son manifestaciones de constructos dentro de una teoría articulada que está bien sustentada en datos empíricos.

Por eso, cuando hacemos validación de constructo lo que hacemos es enunciar una teoría que sugiere variables y sus relaciones y contrastar dicha teoría de forma empírica (NUNNALLY y BERNSTEIN, 1995). De esta forma, la validez de constructo nos indica el grado en que un instrumento de evaluación es una medida adecuada del constructo y en qué nivel las hipótesis derivadas del mismo pueden confirmarse mediante la utilización del instrumento (GARCÍA RAMOS, 1986).

El proceso de validación de un constructo no difiere en su esencia del proceso científico que se utiliza para desarrollar y contrastar teorías (ROSALES, 1997). Las tareas que permitirán la validación del constructo son (CAMACHO y SÁNCHEZ, 1997):

1. Formulación del modelo, así como especificación de las relaciones entre las variables y los constructos (hipótesis).
2. Elaboración de los ítems o indicadores que representen manifestaciones específicas del constructo.
3. Recogida de datos que permita comprobar las hipótesis establecidas.
4. Interpretación y explicación de los datos obtenidos a la luz del modelo propuesto.

Esto implica que la validez de constructo no puede expresarse en la forma de un coeficiente simple, sino que resultan necesarias múltiples investigaciones, además de integrar adecuadamente los resultados de todas ellas (CRONBACH y MEEHL, 1955).

Determinadas técnicas estadísticas facilitan la validación del constructo. Una de las más sencillas, aunque también más débiles, es estudiar las *correlaciones* que se establecen entre las distintas partes de un test. Un ejemplo de esto lo tenemos en un estudio previo en el que analizamos el constructo *preocupaciones en el desarrollo profesional adulto* (PADILLA, 1999). En este caso, partíamos de un instrumento basado en una teoría del desarrollo profesional que contemplaba la existencia de un conjunto de etapas en dicho proceso (exploración, establecimiento, mantenimiento y desenganche), dentro de las cuales existían a su vez distintas sub-etapas. Un indicio de la validez de constructo de la prueba consistía en que las preocupaciones propias de las distintas sub-etapas presentaban una correlación muy alta con las preocupaciones generales de la etapa en que se incardinaban, y más baja con el resto de etapas del desarrollo. La cuestión, por tanto, estriba en comprobar si los resultados de la aplicación de una prueba reproducen la estructura de relaciones que predice el constructo desde el cual está construida.

Una forma similar de estudiar la validez de constructo, aunque mucho más exhaustiva que el procedimiento anterior, consiste en el *análisis factorial exploratorio*. En este caso, comprobamos las correlaciones ítem a ítem para ver si los distintos indicadores observables del constructo (ítems) correlacionan entre sí como predice la teoría. En el caso de que exista correspondencia entre el constructo y sus indicadores, éstos tenderán a formar grupos (formados por ítems que correlacionan de manera alta entre sí y de manera más baja con el resto) que reflejan la estructura teórica del instrumento. El análisis factorial exploratorio es una de las técnicas más empleadas en la validación de constructos.

Un procedimiento aún más riguroso para la validación de constructos es el *análisis factorial confirmatorio*. Es muy parecido al procedimiento anterior. En este caso, y siempre desde el modelo teórico postulado en el instrumento, se exige que ciertas correlaciones entre ítems valgan cero y que otras sean significativamente distintas de cero, mientras que en el análisis factorial exploratorio nos conformamos con que se dé un patrón determinado de correlaciones sin imponer estas restricciones.

Entre las técnicas más sofisticadas para la validación del constructo destacan las *matrices multirrasgo-multimétodo* (CAMPBELL y FISKE, 1959 y 1984) que, básicamente, consisten en medir varios constructos por varios métodos. Si los diferentes métodos son coincidentes a la hora de medir un mismo rasgo, entonces hay ciertas garantías de la solidez del rasgo en cuestión (validez convergente); por otro lado, cuando se miden diferentes rasgos por el mismo método deberá haber divergencia (validez discriminante). Al requerir trabajar con varios constructos y con varias medidas es un procedimiento complejo que se utiliza menos que los anteriores.

Los procedimientos mencionados son los que más se emplean en la validación de constructos, aunque no los únicos. Su presentación en este apartado debe permitir destacar una idea central: que la validez de un instrumento exige poner en práctica diferentes procedimientos para su estimación, en un proceso continuo de acumulación de datos.

Si bien no todos los profesionales del diagnóstico participan activamente en estudios de validación de instrumentos, como usuarios de tests sí es preciso que tengan unos conocimientos mínimos sobre validez y fiabilidad. Estos conocimientos les permitirán adoptar una postura crítica ante el instrumento que se disponen a usar. Es importante que, cuando estos profesionales revisen las guías técnicas de los tests, sepan entender los procesos y resultados que allí se presentan en torno a la estimación de la validez y la fiabilidad del instrumento. De esta forma, podrán percibir de primera mano si los estudios de validación realizados sobre la prueba que se disponen a usar garantizan o no los mínimos necesarios para una adecuada aplicación e interpretación de la información.

Referencias bibliográficas

ADAMS, G. S. (1983). *Medición y evaluación*. Barcelona: Herder.

AMERICAN EDUCATIONAL RESEARCH ASSOCIATION, AMERICAN PSYCHOLOGICAL ASSOCIATION Y NATIONAL COUNCIL ON MEASURMENT IN EDUCATION (1985). *Standars for educational and pyschological testing*. Washington, DC: Amercian Psychological Assocation.

ANASTASI, A. (1982). *Tests psicológicos*. Madrid: Aguilar.

ANASTASI, A. (1992). What counselors should know about the use and interpretation of psychological tests. *Journal of Counseling and Development, 70* (5), 610-615.

ANAYA, D. y REPETTO, E. (1994). Estudio de la madurez vocacional de los estudiantes de educación secundaria mediante el Career Development Inventory. *Educación, 21*, (1), 159-174.

ASENSI, J. y LÁZARO, A. (1979). *Vademécum de pruebas psicopedagógicas*. Madrid: Ministerio de Educación y Ciencias.

BORGATTA, E. F. y BOHRNSTEDT, G. W. (1981). Level of measurement: once over again. En E.F. Borgatta y G.W. Bohrnstedt (eds), *Sociological measurement. Current Issues* (pp. 23-38). Beverly Hills, CA: Sage.

BROWN, F. G. (1980). *Principios de la medición en psicología y educación*. México, DF: Editorial El Manual Moderno, S.A.

CABRERA, F. A. y ESPÍN, J. V. (1986). *Medición y evaluación educativa*. Barcelona: PPU.

CAMACHO, C. y SÁNCHEZ, E.F. (1997). *Psicometría*. Sevilla: Kronos.

CAMPBELL, D. T. y FISKE, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.

CAMPBELL, D. T. y FISKE, D. W. (1984). Las validaciones convergente y discriminante mediante la matriz multirrasgo-multimétodo. En G.F. Summers (coord.), *Medición de actitudes* (pp. 128-152). México, DF: Trillas.

CLARK, L. A. y WATSON, D. (1995). Constructing validity: basic issues in objective scale development. *Psychological Assessment, 7* (3), 309-319.

CROCKER, L. y ALGINA, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston.

CRONBACH, J. L. (1972). *Fundamentos de la exploración psicológica*. Madrid: Biblioteca Nueva.

CRONBACH, L. J. y MEEHL, P. E. (1955). Construct validity in psychological test. *Psychological Bulletin, 52*, 281-302.

DE KETELE, J. M. y ROEGIERS, X. (1995). *Metodología para la recogida de información*. Madrid: La Muralla.

ELLIS, M. V. y BLUSTEIN, D. L. (1991). Developing and using educational and psychological tests and measures: the unifictionist perspective. *Journal of Counseling and Development, 69*, 550-563.

FIGUERA, P. (1996). *La inserción del universitario en el mercado de trabajo*. Barcelona: EUB.

GARCÍA JIMÉNEZ, E. (1997). *La recogida de información*. Material de la asignatura Diagnóstico Pedagógico, inédito. Sevilla: Facultad de Ciencias de la Educación.

GARCÍA JIMÉNEZ, E. y GARCÍA VIDAL, J. (1997). Generando instrumentos de diagnóstico en educación: construcción, uso e interpretación de tests. En H. Salmerón (coord.), *Diagnosticar en educación* (pp.125-161). Granada: FETE-UGT.

GARCÍA, E., GIL, J. y RODRÍGUEZ, G. (1996). *Introducción a la teoría clásica de los tests*. Sevilla: Grupodelta.

GARCÍA RAMOS, J. M. (1986). Validación de constructo en el ámbito pedagógico. *Revista Española de Pedagogía, 147*, 535-554.

GIL, J., GARCÍA, E. y RODRÍGUEZ, G. (1998). Comparación de métodos estadísticos para la evaluación del funcionamiento diferencial de los ítems. *Revista de Investigación Educativa, 16* (1), 187-205.

GONZÁLEZ BLASCO, P. (1986). Medir en las ciencias sociales. En M. García Ferrando, J. Ibáñez y F. Alvira (comp.), *El análisis de la realidad social. Métodos y técnicas de investigación* (pp. 209-267). Madrid: Alianza Editorial.

GORDON, L. V. (1995). *Cuestionario de valores interpersonales*. Madrid: TEA.

GRONLUND, N. E. (1974). *Elaboración de tests de aprovechamiento*. México DF: Trillas.

HALADYNA, T. M. (1994). *Developing and validating multiple-choice test items*. Hillsdale, NJ: Lawrence Erlbaum Associates.

HAMBLETON, R. K. (1980). Contributions to criterio-referenced testing technology. *Applied Psychological Measurement, 4* (4), 421-581.

JORNET, J. M. y SUÁREZ, J. M. (1996). Pruebas estandarizadas y evaluación del rendimiento: usos y características métricas. *Revista de Investigación Educativa, 14* (2), 141-163.

KERLINGER, F. N. (1975). *Investigación del comportamiento*. México DF: Nueva Editorial Interamericana.

LÁZARO, A. (1986). *Orientación y educación especial*. Madrid: Anaya.

LEVINE, S. y ELZEY, F. F. (1973). *Introducción a la medición en psicología y en educación. Manual autoprogramado*. Buenos Aires: Paidós.

LOEVINGER, J. (1957). Objetive tests as instruments of psychological theory. *Psychological reports, 3*, 635-694.

MARTÍNEZ ARIAS, R. (1995). *Psicometría: teoría de los tests psicológicos y educativos*. Madrid: Síntesis.

Mc DIVITT, P. J. y St. JOHN, R. (1996). The use of alternative assessments in adult career conunseling. *Career Planning and Adult Development Journal, 12* (2), 55-72.

MEHRENS, W. A. y LEHMANN, I. J. (1982). *Medición y evaluación en la educación y en la psicología*. México, DF: Compañía Editorial Continental, S.A.

MESSICK, S. (1980). Test validity and the ethics of assessment. *American Psychologist, 35* (11), 1012-1027.

MESSICK, S. (1995). Validity of psychological assessment. Validation of inferences form persons' responses and performances as scientific inquiy into score meaning. *American Psychologist, 50* (9), 741-749.

MORALES, P. (2000). *Medición de actitudes en psicología y educación. Construcción de escalas y problemas metodológicos*. Madrid: Universidad Pontificia Comillas.

MORELAND, K. L. et al. (1995) Assessment of test user qualifications. *American Psychologist, 5* (1), 14-23.

NUNNALLY, J. C. y BERNSTEIN, I. J. (1995). *Teoría psicométrica*. México DF: McGraw-Hill.

PADILLA, M.T. (1999). *El desarrollo profesional en la edad adulta: adaptación y validación del Adult Career Concerns Inventory (ACCI) en la población española*. Tesis doctoral inédita. Sevilla: Facultad de Ciencias de la Educación de la Universidad de Sevilla.

PADILLA, M. T. (2001). *Estrategias para el diagnóstico y la orientación profesional de personas adultas*. Barcelona: Laertes.

PEIRÓ, J. M. et al. (1987a). La medida de la experiencia de la ambigüedad en el desempeño de roles: el cuestionario general de ambigüedad de rol en ambientes organizacionales. *Evaluación Psicológica, 3* (1), 27-53.

PEIRÓ, J. M. et al. (1987b). La medida de la experiencia de conflicto en el desempeño de roles: el cuestionario de conflicto de rol. *Evaluación Psicológica, 3* (3), 345-381.

PEREIRA, M. (1992). *La intervención educativa para el desarrollo de la carrera. Análisis de las necesidades para la carrera de los estudiantes al finalizar la*

educación secundaria. Tesis doctoral inédita. Oviedo: Facultad de Ciencias de la Educación de la Universidad de Oviedo.

PÉREZ JUSTE, R. (1989). Los datos para la evaluación. En R. Pérez Juste y J.M. García Ramos, *Diagnóstico, evaluación y toma de decisiones* (pp. 73-193). Madrid: Rialp.

PÉREZ JUSTE, R. (1991). *Pedagogía experimental. La medida en educación*. Madrid: UNED.

POPHAM, W. L. (1983). *Evaluación basada en criterios*. Madrid: Magisterio Español.

RODRÍGUEZ MORENO, M. L. (1999). *Enseñar a explorar el mundo del trabajo*. Málaga: Aljibe.

ROSALES, G. (1997). La validez de constructo en educación. Alcances y límites. *Ciencia, Ergo Sum, 4* (2). Disponible en http://ergosum.uaemex.mx/julio97/rosales.html.

STEVENS, S. S. (1951). Mathematics, measurement and psychophysics. En S. S. Stevens, *Handbook of experimental psychology*. New York: Wiley and Son.

THORNDIKE, R. L. y HAGEN, E. (1978). *Tests y técnicas de medición en psicología y educación*. México, DF: Trillas.

WOLF, R. M. (1990). The nature of educational evaluation. En H.J. Walberg y G.D. Haertel (eds.), *The international encyclopedia of educational evaluation*. Oxford: Pergamon Press.

CAPÍTULO

5

OTRAS TÉCNICAS Y ESTRATEGIAS PARA EL DIAGNÓSTICO EN EDUCACIÓN

Eduardo GARCÍA JIMÉNEZ, Isabel LÓPEZ GÓRRIZ, Emilia MORENO SÁNCHEZ y Mª Teresa PADILLA CARMONA

1. INTRODUCCIÓN

Este capítulo que nos disponemos a abordar se aparta un poco de la tónica general seguida en los anteriores. En primer lugar, hemos contado con la participación de otros autores que nos han brindado su dilatada experiencia en el estudio y aplicación de algunas técnicas y estrategias de diagnóstico. Así, el profesor Eduardo García ha confeccionado el apartado relativo al portafolios; la profesora Isabel López ha elaborado el material sobre las historias de vida; la profesora Emilia Moreno ha sintetizado, en el apartado correspondiente, su experiencia de trabajo con la técnica de grupo de discusión; y, finalmente, la que suscribe, ha elaborado el apartado dedicado a las rejillas de constructos personales.

En segundo lugar, este capítulo se diferencia de los anteriores también en su enfoque. Aquí no vamos a centrarnos en un conjunto de técnicas que tienen una base común, como sucede, por ejemplo, con las técnicas observacionales, las de encuesta, o las de medición. A diferencia de estas estrategias, entre el portafolios y la historia de vida, y entre éstos y el grupo focal o la rejilla de constructos personales, existen múltiples diferencias y no parece encontrarse un patrón claro que nos permita clasificarlas bajo un único grupo. Por ello, existe una mayor diversidad en el contenido de este capítulo o, dicho de otra forma, se abordan más procedimientos, pero de manera más breve.

El hecho de que hayamos denominado el capítulo como *Otras técnicas y estrategias para el diagnóstico en educación* ya nos indica que existe cierta dificultad para agrupar el conjunto de técnicas que aquí se presentan. Otras denominaciones como *estrategias alternativas de diagnóstico*, *estrategias de autoinforme*, o *técnicas cualitativas*, etc. no serían aplicables a la totalidad de los procedimientos que aquí se incluyen. Por lo que se ha preferido una denominación genérica que, si bien no discrimina sus características peculiares, al menos, nos permite diferenciarlas del resto de técnicas e instrumentos abordados en otros capítulos de esta obra.

Sin embargo, todos los procedimientos que se pueden usar en el proceso diagnóstico no se agotan en las estrategias tradicionales de observación, encuesta o medición. Existen otras formas de recoger y procesar información que aportan nuevas posibilidades al diagnóstico educativo. De hecho, los procedimientos mencionados en este capítulo (portafolios, historia de vida, grupo focal y rejillas de constructos personales) sólo son una pequeña muestra de los que se pueden utilizar, aunque, por cuestiones de espacio, hemos preferido omitir muchos de ellos.

Los criterios que hemos tenido en cuenta para su selección son varios. Por un lado, nos parecía importante incluir aquí instrumentos de cierto impacto en el diagnóstico y la investigación educativa actuales. En este caso, tanto el portafolios como la historia de vida son estrategias que recientemente están recibiendo una considerable atención por

parte de investigadores y profesionales, lo que nos lleva a considerarlas como de obligada referencia en la formación de futuros profesionales del diagnóstico educativo.

Por otro lado, cada una de las estrategias seleccionadas representa una forma distinta de proceder en el diagnóstico. Estos procedimientos diferenciales nos permiten ampliar la clasificación de técnicas diagnósticas implícita en la estructura de esta obra. De esta forma, a los procedimientos de recogida de información basados en la observación, la encuesta y la medición, podríamos añadir tres grupos más de procedimientos:

— Las *técnicas basadas en el grupo*, en las que un grupo de sujetos —y no los sujetos considerados individualmente— actúa como generador y productor de la información. Además del grupo de discusión, bajo este criterio, existen otras posibles técnicas grupales como son las dinámicas de grupo, el grupo *delphi*, etc.

— Las *técnicas basadas en el autoinforme libre y la introspección,* mediante las cuales el sujeto expresa de forma libre informaciones sobre sí mismo y su experiencia. Los autoinformes, por otra parte, son producto de la introspección o la auto-observación que una persona realiza sobre su propia conducta (PÉREZ-PAREJA, 1997). Dentro de este grupo podemos incluir la historia de vida, los ensayos libres autodescriptivos, las técnicas de pensamiento en voz alta... Es necesario hacer constar aquí que los cuestionarios e inventarios de personalidad incluidos en el capítulo anterior también son considerados como técnicas de autoinforme, aunque aquí hemos preferido incluirlos dentro del grupo de los instrumentos de medida.

— Las *técnicas basadas en construcciones y ordenaciones personales*, en las que el sujeto selecciona, construye y ordena la información, generalmente sobre sí mismo o sobre su experiencia vital. Dentro de este grupo podríamos encuadrar al portafolios y las rejillas de constructos personales, además de otros procedimientos como la técnica Q y las tarjetas vocacionales.

De cualquier forma, esta clasificación no es sino una propuesta personal en la que se pretende recoger un conjunto amplio de procedimientos que resulta difícil categorizar dentro de la observación, la encuesta y la medición. En realidad, todas las técnicas e instrumentos de diagnóstico, especialmente los incluidos en este capítulo, no se encuadran con facilidad y plenitud en su correspondiente categoría. Muchas veces una determinada estrategia se asemeja bastante a otra, o combina en su procedimiento varias estrategias distintas de recogida de información. Pensemos por ejemplo en una escala de estimación; en su formato se parece bastante a una escala de actitudes o a un inventario. O pensemos en el grupo de discusión; en su aplicación se genera una auténtica situación de entrevista.

Como éstos se pueden poner múltiples ejemplos que demuestran la dificultad de encasillar o etiquetar a las distintas estrategias de diagnóstico. En lugar de lamentar esta situación, debemos entender esta complejidad como necesaria, si queremos disponer de procedimientos de recogida de información que sean capaces de adaptarse al dinamismo propio de la realidad educativa.

2. EL PORTAFOLIOS

La carpeta o portafolios es un procedimiento de *evaluación de ejecuciones*, que se apoya en la recogida y almacenamiento de información sobre los *logros o adquisiciones* hechos por una persona durante un *período de formación*.

Como los proyectos, los ejercicios, las demostraciones y las observaciones hechas por un supervisor (profesor, jefe de departamento, etc.), los portafolios son una forma de evaluación integrada dentro del modelo de evaluación basada en ejecuciones y vinculada a la evaluación de logros educativos (*outcome-based evaluation*).

Una evaluación basada en ejecuciones (*performance assessment*) requiere que aquellos que están siendo evaluados (alumnos, profesionales, personas que buscan su primer empleo...), en un momento determinado, demuestren, construyan, desarrollen un producto o solución a partir de unas condiciones y estándares definidos. En este sentido, la utilización de los portafolios implica que los alumnos, profesionales o personas que buscan su primer empleo se ven obligados a estructurar las tareas por las que van a ser evaluados, aplicar sus conocimientos previos, elaborar sus respuestas e incluso explicar el proceso que les ha llevado a las mismas (RODRÍGUEZ ESPINAR, 1997:187).

La información recogida en un portafolios se refiere a un *período de tiempo*, equiparable a las evaluaciones escolares o a los niveles de formación continua. De modo que pueden existir portafolios de corto, medio y largo plazo, de curso, nivel, etapa, etc. o, como señalan COLE, RYAN y KICK (1995:11):

a) Los portafolios *de proceso,* que muestran los progresos que van haciendo las personas.

b) Los portafolios *de producto,* que muestran lo aprendido al final de un período de formación.

Del mismo modo, es posible identificar otras modalidades de portafolios considerando dimensiones tales como las tareas a evaluar, la puntuación y valoración, la actividad sobre la que se recoge información o la propia estructura y contenido de los portafolios. Así, las tareas a evaluar pueden situarse en diferentes *momentos del proceso formativo*: podemos encontrar portafolios que se centren en un estadio inicial de la formación mientras que otros lo hacen en el resultado de la misma.

La *puntuación y valoración* de un portafolios puede realizarse desde dos enfoques: uno, que considera cada trabajo por separado; y, otro, que valora todos los trabajos que integran el portafolios. Asimismo, la información recogida en un portafolios puede estar referida a una sola tarea o a *más de un tipo de actividad* diferente.

La *estructura y el contenido* de los portafolios están estrechamente ligados a la finalidad y/o área bajo la que se desarrolla el portafolio. Por consiguiente, existen

también diferentes modalidades de portafolios, según esta dimensión:

a) Portafolios de *enseñanza*: referido a la actuación docente.

b) Portafolios *de aprendizaje*: entendido como una estrategia de aprendizaje.

c) Portafolios *profesionales*: relacionado con los procesos de inserción o reinserción profesional y social.

Por cuestiones de espacio, aquí profundizaremos sólo en los portafolios de aprendizaje, aunque muchas de sus características son también aplicables a los otros dos tipos (portafolios de enseñanza y portafolios profesionales).

2.1. Portafolios de aprendizaje

El portafolios de aprendizaje es un procedimiento para evaluar las ejecuciones de los alumnos, un conjunto sistemático y organizado de evidencias que utilizan alumno y profesor para *controlar el progreso* de las actitudes, las habilidades y el conocimiento. En esencia, un portafolios de aprendizaje se construye para responder a un *propósito evaluativo*, de modo que comparando el contenido del portafolios con determinados criterios sea posible valorar la calidad de las ejecuciones y, sobre todo, ayudar a la *toma de decisiones* sobre los alumnos a nivel individual, de programas, centros y otras variables educativas.

En la definición de los portafolios de aprendizaje, con ser esencial el elemento evaluativo, no debemos olvidar otros componentes que definen su naturaleza:

a) Un procedimiento *metacognitivo*, que ayuda al alumno a comprender y ampliar su propio aprendizaje.

b) Una oportunidad para el lector de *profundizar sobre aprendizaje y aprendiz* (PORTER and CLELAND, 1995:23).

Construir portafolios de aprendizaje que hagan posible un proceso de evaluación alternativa supone recoger, sintetizar e interpretar información de modo que pueda conocerse cómo ha progresado un estudiante o una clase y si está/n preparado/s para afrontar nuevos retos educativos. Contar con una información válida y apropiada es, entonces, tan importante como partir de un esquema comparativo apropiado.

Para que los portafolios cumplan con ese propósito evaluativo se requiere un esquema de trabajo bien elaborado y una reflexión continua sobre cuestiones como:

¿Cuál es el propósito de la evaluación?

¿Qué tareas deberían incluirse en el portafolios?

¿Qué estándares o criterios habría que aplicar?

¿En qué medida podría asegurarse la consistencia de las puntuaciones o valoraciones emitidas?

¿Responden los resultados a los propósitos perseguidos con el portafolios?

¿Cómo se han utilizado los resultados? (HERMAN, GEARHART, ASCHBACHER, 1996:29)

Tomemos estas cuestiones como referencia básica para presentar un posible diseño de portafolios de aprendizaje. Así, consideramos que el diseño de los portafolios de aprendizaje debería contemplar los siguientes elementos:

- Propósitos del portafolios de aprendizaje, es decir, para qué se va a elaborar dicho portafolios.
- Contenidos del portafolios, que deben ser estructurados para reflejar tanto los propósitos evaluativos del portafolios como las metas de aprendizaje que se pretende que alcance el alumno. Por tanto, en un portafolios deberían incluirse:

  a) Las metas de aprendizaje.

  b) Las tareas de aprendizaje.

  c) Un modelo que estructure los contenidos.

  d) Los materiales o artefactos.

  e) Los criterios de puntuación.

  f) La validez del portafolios.

2.2. Propósitos de los portafolios de aprendizaje

Artistas, arquitectos, fotógrafos, modelos y otros profesionales noveles se apoyan en las carpetas o portafolios para *demostrar sus aptitudes y sus logros*. Dentro de la carpeta incluyen muestras de su trabajo que *ejemplifican* la amplitud y profundidad de sus aptitudes. Con unas fuentes de información tan ricas es más fácil para críticos o enseñantes, y desde luego para los propios interesados, *comprender* el desarrollo de las aptitudes y *planificar* las experiencias que estimulen progresos adicionales y logros a diferentes niveles: certificaciones, exposiciones, mercado, empresas, etc. El trabajo realizado a partir del uso de los portafolios de aprendizaje parte de la postura de que los estudiantes de todos los niveles del sistema educativo y de la educación no formal no merecen menos.

Los portafolios de aprendizaje son útiles a diferentes propósitos relacionados con la valoración y el progreso de los alumnos y afectan, sin duda, a la labor de los profesores y de otros miembros de la comunidad educativa. En el cuadro 1 revisamos las principales posibilidades asociadas al uso de portafolios de aprendizaje.

CUADRO 1

POSIBILIDADES ASOCIADAS AL USO DEL PORTAFOLIOS DE APRENDIZAJE

DIMENSIONES	PROPÓSITOS
Evaluación del currículo	• Evaluación externa. • Evaluar el programa o la adecuación del currículo.
Certificación individual de logros	• Evaluar el progreso individual del estudiante en diferentes áreas. • Determinar el progreso a través de los ciclos y niveles del sistema educativo. • Certificar oficialmente los logros: p.e. Graduado en Enseñanza Secundaria o Diploma de estudios CCC.
Toma de decisiones	• Diagnosticar las necesidades del estudiante. • Aportar información para planificar la enseñanza en la clase. • Mejorar la calidad de la enseñanza.
Desarrollo profesional	• Fomentar la eficacia docente. • Estimular la práctica reflexiva a nivel de aula y de centro. • Intercambiar puntos de vista con otros profesores y profesionales de la educación. • Apoyar el desarrollo profesional de los profesores.
Evaluación alternativa	• Fomentar la eficacia del estudiante. • Promover la autoevaluación del alumno. • Hacer reflexionar al estudiante sobre los cambios y el desarrollo operado a lo largo del tiempo. • Fomentar en el estudiante la responsabilidad hacia su propio aprendizaje. • Hacer reflexionar al estudiante sobre el currículo aprendido. • Estimular el trabajo del estudiante. • Hacer reflexionar al estudiante sobre las estrategias que apoyen su aprendizaje. • Hacer reflexionar al estudiante sobre las metas que deben presidir futuras experiencias.
Audiencias	• Establecer vías de comunicación con los padres.

2.3. Las metas de aprendizaje

Tomando como referencia básica lo apuntado en el diseño del portafolios de aprendizaje, el primer aspecto que debe considerarse es el de las metas que orientan las realizaciones del estudiante. Puesto que la decisión sobre lo que se debe evaluar tiene que estar basada en prioridades curriculares y educativas, el paso crucial de esta fase es determinar las metas fundamentales de la enseñanza en tanto que centro educativo, nivel de enseñanza, área de conocimientos, equipo docente, etc.

Las metas válidas para orientar las actuaciones de los alumnos serán aquellas que:

a) Reflejen de un modo prioritario los resultados del estudiante.

b) Hagan referencia a contenidos disciplinares de carácter fundamental.

c) Lleven al estudiante a una comprensión profunda de lo aprendido, le exijan un pensamiento complejo y la resolución de problemas.

d) Sean significativas para el estudiante.

e) Resulten adecuadas a su nivel de desarrollo y ritmo de aprendizaje.

Deben proponerse metas u objetivos amplios y no aptitudes aisladas demasiado específicas u objetivos relacionados con temas o lecciones concretas. Como es lógico, esas metas difieren entre sí en función de los currículos considerados en cada caso. En el cuadro 2, presentamos un ejemplo de Metas relativas al Aprendizaje de la Escritura.

2.4. Contenidos del portafolios

Siguiendo el diseño del portafolios de aprendizaje, los contenidos que en él deben incluirse están condicionados por las metas bajo las que se estructuran las realizaciones del estudiante.

Resulta atrevido, pues, proponer una relación única de contenidos. No obstante, en términos generales, podría hacerse una propuesta que recogiese los siguientes elementos:

a) Un objetivo claramente identificado.

b) Fragmentos de trabajos seleccionados por el alumno.

c) Muestras de valoraciones formativas y sumativas realizadas por él y por otras personas.

d) Reflexiones del alumno.

e) Muestras que ejemplifiquen el progreso del estudiante.

Respecto a cómo deben colocarse los contenidos en los portafolios, es conveniente organizarlos en dos niveles complementarios que incluirían:

a) Evidencia real o datos brutos, que se incluyen en la carpeta. Permite a los profesores examinar el contenido real y las notas sobre el avance de los alumnos.

b) Resumen o marco de organización que ayude a sintetizar la información anterior. Facilita a los profesores la toma de decisiones y la comunicación con los padres u otros profesores y profesionales de la enseñanza.

CUADRO 2

EJEMPLO DE METAS RELATIVAS AL APRENDIZAJE DE LA ESCRITURA

1. ¿Qué principios y conceptos, que resultan fundamentales para un conocimiento aplicado de esta disciplina, me gustaría que comprendiesen y utilizasen mis estudiantes?

1. Comprender cómo están estructurados los diferentes géneros narrativos (utilizar los componentes temáticos, contexto, representación, uso del lenguaje, etc.) y los modos en que han sido adaptados por los autores más representativos; elaborar composiciones que estén apoyadas en los géneros estudiados.
2. Analizar acontecimientos de la actualidad desde las perspectivas histórica, política, geográfica y multicultural e integrar esas aportaciones en una composición escrita de carácter argumentativo.
3. Comprender las relaciones causa-efecto en la historia y en la vida cotidiana y escribir ensayos integrando argumentos que se apoyen en tales relaciones.
4. Comunicarse tanto oralmente como por escrito.
5. Utilizar diferentes medios para expresar lo que se conoce o experimenta.

2. ¿Qué actitudes y habilidades sociales quiero que desarrollen mis estudiantes?

1. Trabajar de forma independiente.
2. Desarrollar un espíritu de trabajo en equipo y poner en práctica habilidades que faciliten el trabajo en grupo.
3. Apreciar sus puntos fuertes a nivel individual, tener confianza en sus habilidades.
4. Ser persistente ante los desafíos.
5. Estar orgulloso de su trabajo.
6. Valorar el aprendizaje y pasárselo bien.
7. Mantener un sano escepticismo sobre el valor de ciertas afirmaciones y argumentaciones.

3. ¿Qué estrategias analíticas y reflexivas me gustaría que desarrollasen mis estudiantes?

1. Discutir los diferentes modos en que pueden mejorar los planes y estrategias utilizadas para crear proyectos.
2. Reflejar el proceso utilizado para escribir, evaluar su eficacia y establecer planes sobre cómo mejorarlo en el futuro.
3. Elaborar planes eficaces para construir proyectos de forma independiente y monitorizar su progreso hasta llevarlos a término.

(Tomado de HERMAN, GEARHART y ASCHBACHER, 1996: 32)

2.5. Las tareas de aprendizaje

Una tarea de aprendizaje es, en esencia, aquello que tienen que hacer los alumnos para demostrar que han logrado determinados objetivos de aprendizaje. En realidad, las tareas de aprendizaje muchas veces simplemente se derivan de los propios objetivos o metas de aprendizaje. Por ejemplo, si un resultado previsto es:

— Escribir un ensayo que relacione la caída del muro de Berlín con el resultado de las dos guerras mundiales.

— O bien analizar las consecuencias para Alemania del Tratado de Versalles.

En tales metas, básicamente ya están presentes los rasgos generales de las posibles tareas a realizar.

Las buenas tareas para incluir en un portafolios son, precisamente, las que están relacionadas con las buenas estrategias de enseñanza. Es decir, aquellas estrategias que invitan a los alumnos a que aborden contenidos verdaderamente significativos —temas, principios y conceptos fundamentales de una materia— y les piden que construyan sus propias visiones de las cosas y que apliquen lo que han aprendido a problemas de cierta complejidad pero que despierten su interés.

Dado que en cualquier propuesta curricular existen muchos contenidos de interés sobre los que pueden plantearse cometidos que merecen la atención de alumnos y profesores, podrían proponerse tareas que sobrepasen los límites de una asignatura o área de conocimientos y que incorporen metas interdisciplinares de aprendizaje.

Si una meta de aprendizaje fuese valorar el progreso del alumno en la composición escrita, dicha composición podría abordar un tema relacionado con cualquier material del currículo. Un portafolios de escritura podría incluir muestras de composiciones realizadas en ciencias, historia o matemáticas; incluso podría explorar temas transversales —diversidad, cambio, conflicto— que fomentasen conexiones interdisciplinares.

Un tipo fundamental de tareas que puede abordarse en el portafolios es aquella que proporciona al profesor evidencia complementaria sobre la competencia del alumno, pidiendo a éste que reflexione sobre su propio aprendizaje. Se trata de una autoevaluación reflexiva del tipo: «¿qué he aprendido sobre...?».

2.6. Modelos de portafolios de aprendizaje

En esencia, entre los diferentes modelos de portafolios de aprendizaje que pueden desarrollarse el más adecuado es aquel que mejor responde a los propósitos y las metas de aprendizaje. Las siguientes cuestiones pueden orientarnos acerca del modelo de portafolios más adecuado en cada caso:

- *¿Quién decide cuál es el contenido?*
  - — El profesor.
  - — El alumno.
  - — El alumno asesorado por el profesor o por los demás estudiantes.
- *¿Cómo debería ser el portafolios?*
  - — Un portafolios de exposición (incluiría sólo los mejores materiales).
  - — Un portafolios de progresos (comprendería materiales que mostrasen el progreso del estudiante a lo largo del tiempo).
  - — Un portafolios de trabajo (recogería todos los materiales elaborados por el alumno).
  - — Un portafolios mixto (incluiría una representación de todos los demás tipos de portafolios).
- *¿Qué versión del trabajo debería mostrar?*
  - — La versión final del trabajo del alumno.
  - — Una que incluyese notas, borradores y otras evidencias que muestren el tipo de procesos seguidos por el alumno.
- *¿Qué secciones deberían establecerse?*
  - — Aquellas que mostrasen el trabajo realizado en diferentes momentos del año (evaluaciones).
  - — Unas que se correspondan con las metas generales del currículo seguido.
- *¿Qué materiales deberían incluirse?*
  - — Sólo aquellos que muestren el nivel de competencia logrado por el alumno.
  - — Además de los anteriores, una tabla de contenidos.
  - — Ídem más una visión del propósito del portafolios.
  - — Ídem más información que muestre los antecedentes del estudiante.
- *¿Sobre qué propósitos deberían versar las reflexiones del alumno?*
  - — Progreso.
  - — Puntos fuertes.
  - — Lo aprendido.
  - — Lo que quería o lo que necesitaba aprender.
  - — Sobre el portafolios en su conjunto o sobre alguna parte del mismo.

- *¿Debería preverse un espacio para la reflexión de los padres?*

  — Si así fuese, ¿podrían implicarse en el proceso de evaluación?

  — Sobre qué aspectos podrían opinar, ¿quién debería decidirlo?[26]

A continuación, vamos a mostrar diferentes modelos de portafolios; el primero de ellos (ver cuadro 3) fue desarrollado para los niveles de enseñanza primaria y unitaria de las escuelas de Ohio[27]. Esta propuesta fue desarrollada por estudiantes universitarios de la Junior High (tercer curso) y en alguno de los elementos se pide a los niños que valoren el trabajo que realizan en la escuela.

CUADRO 3

PROPUESTA DE PORTAFOLIOS PARA LOS NIVELES DE LA ENSEÑANZA PRIMARIA Y UNITARIA DE LAS ESCUELAS DE OHIO

NIVELES DE ENSEÑANZA PRIMARIA Y UNITARIA
Southwest Ohio Schools

I. Introducción

A. Página de título (diseñada por el alumno).
B. Declaración del propósito.
C. Reacción del estudiante (septiembre y junio).
D. Renuncia de confidencialidad (incluyendo hoja de forma para los revisores).

II. Quién soy

A. Inventario de intereses.
B. Pensamientos acerca de la *Junior High*.
C. Hábitos de trabajo (inventario, habilidades sociales, personales).
D. Habilidades de toma de decisiones (ejercicios para desarrollar habilidades con las que enfrentarse a la agresividad, la resolución de conflictos, rechazo a participar en actividades dañinas).
E. Actividades de autoconcepto (*Coopersmith Self-Esteem Inventory*).

III. Establecimiento y revisión de metas

A. Académicas.
B. Comportamentales.
C. Relativas a la asistencia.
D. Extracurriculares (sociales).
E. Vitales.

IV. Aspectos académicos y reflexiones

A. Lenguaje (predominante).
B. Ciencias (mejores informes de laboratorio, tests).
C. Ciencias sociales (proyectos, itinerarios de viajes, tests).
D. Salud/Prevención (proyectos al servicio del aprendizaje).
E. Matemáticas (proyectos, tests, tests de ubicación, cálculo).
F. Música (solos o en coro, cosas relacionadas con conciertos).
G. Arte (láminas de proyectos, evaluaciones, etc.)

El siguiente modelo (ver cuadro 4) está pensado también para alumnos de enseñanza primaria (grado 5); asimismo, fue desarrollado en una escuela del estado de Ohio[28]. Resulta un modelo menos académico y que abre de un modo explícito la posibilidad al alumno para que comente aspectos personales no vinculados directamente al currículo escolar.

CUADRO 4

PROPUESTA DE PORTAFOLIOS PARA LOS ALUMNOS DE ENSEÑANZA PRIMARIA DE UNA ESCUELA DE OHIO

ENSEÑANZA PRIMARIA, Grado 5
West Milton, Ohio

I. Página de título e índice

II. Quién soy

A. Presentación.
B. Familia.
C. Amigos.

III. Nivel de éxito escolar

A. Asistencia.
B. Comportamiento.
C. Metas.
D. Habilidades de estudio.
E. Estilo de aprendizaje.
F. Aprendizaje cooperativo.

IV. Éxito académico

A. Lectura.
B. Escritura.
C. Ciencia y Salud.
D. Ciencias Sociales.
E. Matemáticas.
F. Trabajo diario en la clase.
G. Registro de lo escuchado y hablado en la clase.
H. Proyectos especiales.
I. Registros de competencias.

V. Mirando hacia atrás y hacia adelante, reflexiones acerca del curso que se ha terminado.

El tercer modelo que presentamos (cuadro 5) está referido al trabajo en una clase concreta —la de lengua inglesa— y está concebido para recoger las aportaciones de los alumnos en relación con dicha materia. Es, por tanto, un modelo más específico que hace posible la vinculación de este procedimiento valorativo a una materia concreta del currículo escolar[29].

CUADRO 5

PROPUESTA DE PORTAFOLIOS PARA LA CLASE DE LENGUA INGLESA

Senior English Class

***Propósito*:** Autoevaluación de habilidades y progreso en la escritura de modelos de respuesta lectura /escritura.

I. Escritura

A. Elección de tres ensayos.
   1. Respuesta a cuestiones relacionadas con la Literatura.
   2. Ensayo personal.
   3. Trabajo que implique la defensa de una idea o argumento.
   4. Trabajo que analice una relación causa-efecto.
   5. Trabajo comparativo.
   6. Trabajo que realice alguna clasificación.

B. Forma paralela editada para cada trabajo.
C. Un trabajo de investigación.
D. Reflexión sobre el trabajo de investigación.
E. Elección de una pregunta tipo ensayo sobre literatura (a partir de una lista).
F. Reflexiones que justifiquen por qué se ha elegido cada trabajo y cada pregunta.

II. Lectura

A. Reflexiones sobre cualquier pieza literaria (lista).
B. Reflexiones sobre por qué se eligieron las RDG y las respuestas.

III. Resumen

A. Autoevaluación del progreso en escritura.
B. Análisis de la forma paralela, sugerencias seguidas /desechadas con sus correspondientes explicaciones.
C. Análisis de fortalezas y debilidades en la escritura.
D. Evaluación de las respuestas RDG (temas /modelos).

IV. Orientaciones

A. Costo: mínimo (carpetas / cajas).
B. Almacenaje: armario cerrado.
C. Evaluación: puntos por trabajar en PF. Puntos correspondientes a la evaluación del trabajo.
D. Administración: copia para el director /supervisor (tutor).

El último de los modelos que presentamos (cuadro 6) está concebido para favorecer el uso de los portafolios entre los estudiantes universitarios. Su novedad fundamental estriba en la incorporación del plan de carrera como parte del portafolios[30].

CUADRO 6

PROPUESTA DE PORTAFOLIOS PARA ESTUDIANTES UNIVERSITARIOS

MODELO ALTERNATIVO DE HIGH SCHOOL, Grados 10-12
Lima City Schools, Ohio

I. Introducción

A. Página de título.
B. Pre-reflexiones sobre el porqué y la importancia de hacer un diario.
C. Post-reflexiones sobre el porqué y la importancia de hacer un diario.

II. Retrato del estudiante

A. Meta de asistencia 90%.
B. Tópicos para registrar diariamente.
C. Estimación de actitud (104) (ademanes, laboriosidad, cumple con las asignaciones de tiempo).

III. Programa Educativo

A. Resultados del test de habilidad, de la encuesta sobre la droga y del test de autoestima.
B. Formularios de admisión, resultados del test de rendimiento en habilidades básicas (TABE).
C. Metas del alumno.
D. Copias de trabajos excepcionales.

IV. Plan de carrera

A. Plan educativo y de carrera.
B. Metas a corto y largo plazo.

V. Conclusión

A. Reflexión-metas para volver a introducir dentro del calendario regular de clases.
B. Evaluar las metas a corto plazo.
C. Expectativas para el futuro.

2.7. Materiales de los portafolios de aprendizaje

Los portafolios están concebidos como una herramienta abierta a la incorporación de materiales y recursos de diferente naturaleza. Así, cuando se elaboran portafolios es

posible incluir en ellos materiales que indican estrategias o modos diferentes de presentar la información.

En un portafolios, entonces, pueden recogerse entradas de diarios (pensamientos sobre lo leído o trabajado en clase), anotaciones sobre un libro u otro material de lectura, conversaciones escritas que representan ideas compartidas con compañeros, conversaciones mantenidas con alumnos de otras clases (diario viajero), primeras ideas creativas sobre lo leído o trabajado en clase (esbozos), reflexiones sobre lo debatido en clase, comentarios del círculo de autores de los portafolios (revisiones y recomendaciones de compañeros sobre el trabajo que se está elaborando, sobre todo en ensayos y composiciones escritas), reflexiones sobre el proceso de aprendizaje o una visión del proceso utilizando fotografías.

Los portafolios también incluyen materiales que representan aprendizajes importantes realizados por los alumnos. Así, el recorrido por el trabajo realizado a través de los borradores, materiales relacionados con las celebraciones realizadas en los centros educativos (fotografías, vídeos, dibujos), portadas de los libros utilizados, grabaciones en vídeo y fotografías sobre distintos acontecimientos realizados en el aula o en las instalaciones del centro, trabajos artísticos (esbozos, láminas, poemas, cuentos, representaciones teatrales, etc.), copias de cartas escritas o recibidas de otros alumnos, anotaciones, resultados de pruebas escritas, proyectos, trabajos realizados utilizando el ordenador (programas, bases de datos, trabajos de edición, etc.) o trabajos en grupo recogiendo las valoraciones de los compañeros.

Por último, en los portafolios también es posible recoger otros materiales que en sí mismos son importantes para el propio alumno. Nos referimos al trabajo o proyecto con el que el alumno se siente identificado, los tópicos, ideas o temas que le han gustado, un material que representa una experiencia memorable para el alumno o el material que supone su experiencia más gratificante. Asimismo, el alumno podría incluir el trabajo, proyecto o lectura que ha supuesto un cambio más significativo en su desarrollo académico o personal, el trabajo que mejor expresa sus emociones o que, en algún sentido, resulta único, un nuevo enfoque utilizado al escribir, hacer un trabajo o resolver un problema, el material que mejor refleja los logros del alumno o que muestra cómo ha progresado éste y los cambios que ha realizado entre la primera y la última versión de un trabajo, proyecto, etc. (PORTER y CLELAND, 1995).

2.8. Criterios de evaluación

Fijar los criterios de puntuación es una parte esencial del proceso de evaluación; sin ellos, un portafolios no es más que una colección de trabajos del alumno. Con los criterios se hace público lo que se espera de los alumnos, de modo que los profesores, los padres y los propios alumnos pueden estar en sintonía y cumplir con sus roles respectivos en relación con la mejora del rendimiento del estudiante. HERMAN,

ASCHBACHER y WINTERS (1996) sugieren que los criterios deberían tener las siguientes características:

a) Ser sensibles a la enseñanza y a los propósitos de la evaluación.

b) Estar centrados en lo importante, progresivamente apropiados a los resultados del alumno que reflejan las concepciones cotidianas de la excelencia.

c) Ser significativos y creíbles para los profesores, padres y alumnos.

d) Estar comunicados con claridad.

e) Estar bien elaborados y no sesgados.

Esta lista de características deseables puede utilizarse para orientar las decisiones que es necesario adoptar cuando se elaboran criterios para evaluar los portafolios de aprendizaje. Estas decisiones comienzan a un macro nivel (¿cuál es el foco?) para continuar con cuestiones situadas en un micro nivel, tales como ¿cuántos puntos debería tener la escala?

Los criterios de puntuación deberían estar enfocados sobre aquellos aspectos de la actividad del estudiante que están en consonancia con los propósitos de la evaluación. Dicho de otro modo, la elaboración de los criterios de evaluación debería comenzar respondiendo a tres cuestiones claves:

1. ¿Qué tipo de actividad o rendimiento se está intentando medir?
2. ¿Qué dominios del conocimiento y las habilidades de los estudiantes se desea medir?
3. ¿Quién/es debería/n diseñar los criterios?

Junto a estas cuestiones clave, la elaboración de los criterios de evaluación puede detenerse también en otros elementos de interés:

- La concepción que se tiene sobre la excelencia, es decir, cómo sería la mejor ejecución posible del dominio o tarea que es objeto de evaluación.
- Las puntuaciones, ¿deberían ser holísticas o analíticas?
- Debemos definir y comunicar diferentes niveles de rendimiento; ¿cómo hacerlo?, ¿utilizando los términos al uso por la administración educativa?

2.9. Validez de los portafolios

Diremos que un portafolios de aprendizaje es válido, como procedimiento de diagnóstico-evaluación, cuando proporciona la información que se necesita para tomar buenas decisiones. Por supuesto, esas decisiones están en función del propósito que en cada caso cumpla el portafolios como procedimiento de evaluación.

Por su parte, la validación es un proceso continuo de recogida de evidencias para determinar hasta qué punto es precisa la información que, como procedimiento de evaluación, proporciona el portafolios y cuán apropiado es el uso de esa información. Tampoco debemos olvidar que el portafolios como procedimiento de evaluación es una más de las fuentes de información con las que cuentan los profesores sobre las actuaciones de los alumnos y tan sólo uno de los muchos indicadores que se utilizan a efectos de la promoción o no del alumno y de la planificación de la enseñanza.

De acuerdo con lo apuntado, para aumentar su validez los profesores deberían tener en cuenta algunas de las cuestiones siguientes al diseñar y utilizar el portafolios de aprendizaje (HERMAN, GEARHART and ASCHBACBER, 1996).

1. Las puntuaciones asignadas a los alumnos a partir del portafolios de aprendizaje, ¿reflejan adecuadamente lo que éstos han aprendido?

   El contenido de la evaluación basada en el portafolios, ¿representa las prioridades del profesor sobre el currículo y la enseñanza?

   Los criterios de puntuación, ¿reflejan cómo el profesor percibe los diferentes modos que tienen los alumnos de entender cuáles son las prioridades curriculares?

   Los estudiantes, ¿realmente tienen posibilidad de aprender las cosas que se valoran en el portafolios?

2. Las puntuaciones otorgadas a partir del portafolios de aprendizaje, ¿son lo suficientemente válidas como para hacer generalizaciones acerca del rendimiento de los estudiantes?

   ¿Hasta qué punto cree el profesor que los resultados obtenidos por un alumno representan lo que éste es capaz de hacer en relación con un dominio amplio y significativo del conocimiento?

3. ¿Tienen sesgo las puntuaciones dadas a partir del portafolios de aprendizaje? Lo que los estudiantes han aprendido y lo que han hecho todos los estudiantes, ¿está bien reflejado en las puntuaciones?

   ¿Cuentan los estudiantes con los antecedentes (conocimientos y habilidades previas) necesarios para implicarse con acierto en las tareas del portafolios?

   El contexto y el contenido de las tareas, ¿favorece o perjudica a determinados alumnos?

   Con la evaluación realizada a partir del portafolios, ¿todos los estudiantes pueden demostrar sus habilidades?

4. ¿Existen pruebas que corroboren que la evaluación realizada a partir del portafolios está en la base de una toma de decisiones adecuada?

5. La evaluación basada en el portafolios, ¿contribuye a un aprendizaje y a una enseñanza significativos?

Todas estas cuestiones facilitarán tanto el diseño, como la utilización y la evaluación del portafolios de aprendizaje.

3. LA HISTORIA DE VIDA

En los últimos años se están produciendo importantes cambios tecnológicos, económicos, científicos, sociológicos, educativos, etc., que están afectando y transformando los procesos de arraigo, desarraigo, aculturación, socialización y educación de las personas y de las instituciones.

Estas transformaciones generan nuevas situaciones que nos someten a procesos de crisis personales y sociales, así como a la búsqueda de soluciones de reconstrucción y equilibrio. Esto confronta a las distintas ciencias tecnológicas, humanas, educativas, psicológicas y sociales a buscar recursos, medios e instrumentos que detecten las dificultades y el desequilibrio que sufren las personas y las estructuras sociales, y que ayuden a dar respuestas interventoras restauradoras de equilibrio en todas ellas.

En este sentido, en distintos campos y, en concreto, en los de las ciencias antropológicas, sociales y educativas, se están buscando fórmulas para llegar a las personas e instituciones de manera que se pueda recoger información significativa, que revele los procesos internos que se están generando en ellas, así como sus manifestaciones sintomáticas para hacer una intervención que restablezca un equilibrio evolutivo.

De ahí que se empiecen a utilizar ciertos instrumentos abiertos como son las entrevistas, memorándums, vídeos, historias de vida, etc. Aquí nos vamos a ocupar específicamente de este último, por lo oportuno que nos parece para el campo del diagnóstico y la intervención educativa.

3.1. Emergencia histórica de la historia de vida en sus modalidades biográfica, autobiográfica y de relatos de vida

La historia de vida ha seguido distintos derroteros a lo largo de la historia y ha sido utilizada en distintos momentos por distintas ciencias. Sin embargo, aquí nos limitaremos a situar brevemente su trayectoria y su situación actual.

BERTAUX (1989) hace un breve recorrido histórico de la historia de vida en cualquiera de sus tres formas (biográfica, autobiográfica y de relatos de vida), que nosotros recogemos aquí. Según este autor:

> «La recogida de relatos de vida no es una práctica inocente, no lo dudéis. Contar su vida es poner en juego la imagen de sí, la imagen que los otros se hacen de sí, la imagen de sí para sí mismo. Es correr el riesgo de evocar recuerdos dolorosos ocultos, debilidades que se prefieren olvidar y herencias que uno preferiría guardar secretas. No es un gesto que funcione sin esfuerzo. Y, sin embargo, existen en la cultura occidental en este final de siglo varios elementos que hacen que este gesto sea posible (...). De tal forma que hacer la historia de una vida puede tomar las formas de biografía, autobiografía o de relato de vida» (BERTAUX, 1989:18).

BERTAUX (1989) alude a la *biografía de un gran hombre* como la forma más antigua de historia de vida. En esta forma destacan las biografías de los fundadores de las grandes religiones (Moisés, Cristo, Buda, Mahoma...); las biografías de grandes reyes y conquistadores; y las biografías de santos, cuyas vidas ejemplares servían de modelo a los cristianos.

Esta tradición se ha ido manteniendo y ha ido evolucionando a lo largo del tiempo, así, en los países anglosajones, la biografía de grandes hombres de estado, escritores y artistas se ha convertido en un género literario.

Esta humanización de la biografía aparece como consecuencia de la desaparición de las teocracias y de los absolutismos. Pero esta humanización le debe también bastante a una segunda tradición que es la de la *autobiografía*. San Agustín con sus *Confesiones* fue quien abrió esta vía y Rousseau quien le dio forma moderna. La autobiografía de san Agustín fue una gran novedad en su tiempo, puesto que en la Edad Media se vivía siempre en grupo y era impensable que una vida pudiera ser algo singular.

A partir del siglo XVI, cuando aparecen los burgueses con su sentimiento de identificación con el trabajo, es cuando toma más fuerza la noción de individuo. El Protestantismo refuerza el sentimiento de autonomía de la persona como ser individual, capaz de entrar en diálogo con Dios. Con la emigración de protestantes al Nuevo Mundo, se desarrolla este espíritu individualista. Y el Individualismo, como tal, va ligado a la Revolución americana y a la Revolución francesa, en donde los ciudadanos se convierten en ciudadanos de la república.

El principal efecto cultural de la filosofía individualista es el de transformar y crear la relación de la persona consigo misma, al contrario que las sociedades tradicionales en donde las relaciones de existencia se apoyaban en el grupo. Solamente en la relación del ser consigo mismo es cuando se empieza a pensar en la propia vida como producto de los propios actos, y es así como nace *la autobiografía*.

Esta idea americana, que crea la democracia apoyada en el Individualismo, tardará tiempo en difundirse por Europa. Sin embargo, el siglo XIX será el de la *democratización de la idea autobiográfica*.

En los últimos treinta años, distintos movimientos sociales —todo el movimiento relacionado con la evolución de los roles sexuales y las luchas de género, que ha producido profundas transformaciones familiares— encuentran su fundamento en este concepto de individualismo. Este concepto del individualismo postmodernista confronta a la persona consigo misma, y de ahí que se empiecen a desarrollar movimientos como el psicoanálisis, que permiten recuperar el gran potencial de energía del ser humano para desarrollarse en equilibrio; movimientos de formación desde una perspectiva de *transformación de sí mismo y de autoformación,* apoyándose en la autobiografía (PINEAU y MARIE-MICHÈLE, 1983).

Otra manera en que se ha ido desarrollando en distintos campos (educativos, históricos, sociológicos, antropológicos...) el trabajo en relación consigo mismo, ha sido el *relato de vida*, entendido como «el relato o narración que una persona hace de su experiencia de vida a otra, en una interacción de un cara a cara» (BERTAUX, 1989:28).

En ocasiones, el relato de vida es considerado como una variante de la autobiografía. Sin embargo, y pese a que presentan una forma muy similar, el relato de vida y la autobiografía guardan entre sí diferencias fundamentales. Básicamente la autobiografía (en inglés, *life story*,o *récit de vie*,en francés) corresponde a la historia de una vida tal como la relata la persona que la ha vivido, mientras que el relato de vida (*life history* o *histoire de vie*) alude al estudio de caso referido a una persona dada (ANGUERA, 1995).

BERTAUX (1989) también se refiere a las diferencias entre ambos procedimientos y las sintetiza como exponemos a continuación:

> «El relato de vida ha sido desde el principio la narración autobiográfica del Otro. (...) Al contrario que la biografía, que es espontánea, el relato de vida se inicia por una demanda venida del exterior a alguien que nunca ha tenido la idea de escribir su autobiografía. Esto crea inicialmente una relación de centro-periferia, en donde el investigador, representante de la cultura dominante, el centro, interroga a alguien de la periferia para saber de dónde viene, quién es, qué tipo de cultura le ha hecho tal como es, cómo la ha vivido, cuáles han sido sus experiencias... Escuchando su relato y pidiéndole clarificar lo que para él es normal, el investigador tiende a definir al Otro, como informador de cultura de origen, a orientar su mirada hacia el entorno social en el seno del cual él ha vivido, y que es inaccesible al investigador. El Otro se convierte, pues, en un telescopio, un periscopio, un instrumento de observación» (BERTAUX, 1989: 28-33).

En el relato de vida, a través de la experiencia de la persona interrogada, el investigador no sólo percibe su estructura social, sino que su cultura se hace verbo, se hace oir, se hace visible. De este modo, a través del investigador, la cultura del otro, desde su palabra individual, se convierte en *palabra pública*. A través de esta palabra pública se hacen visibles las relaciones sociales que le dan forma a nuestra existencia, y que suelen aparecer en la cotidianidad como invisibles. Así pues, los relatos de vida permiten ver lo social bajo sus múltiples facetas, no como reflejos desencarnados de las estructuras abstractas, sino como un conjunto articulado de experiencias vividas, que reflejan las lógicas de las estructuras sociales en donde nos ubicamos; estructuras que a su vez nos están modelando.

A través de esta breve mirada histórica, podemos pues decir que la *historia de vida* a lo largo de la historia aparece plasmada bajo tres modalidades diferentes como son: *la biografía, la autobiografía y el relato o narración de vida.* Y, aunque cada una va apareciendo en distintos momentos históricos con finalidades específicas, actualmente todas y cada una de las tres modalidades se utilizan con finalidades literarias, políticas e investigadoras específicas en los distintos campos del saber (literatura, sociología, antropología, medicina, psiquiatría, psicología, educación, etc.).

Tienen una utilización específica y significativa en el campo del *diagnóstico*, ya sea en el diagnóstico clínico —médico, psiquiátrico, psicológico, educativo—, o en un diagnóstico más sociológico o institucional. En este trabajo nos situaremos más concretamente en el diagnóstico educativo-clínico.

3.2. La historia de vida como instrumento y método de diagnóstico e intervención educativa. Presupuestos fundamentales

Desde hace varios años la historia de vida se viene utilizando en el campo de la educación en las distintas ramas que conciernen la formación y autoformación de adultos (PINEAU y MARIE-MICHÈLE, 1983; CENTRO DE ADULTOS «JUAN XXIII», 1999; HORMIGO GONZÁLEZ, 2000); la formación de formadores, formación y orientación profesional, formación de empresa, intervención social, intervención institucional (PINEAU y JOBERT, 1989); la evaluación y la intervención educativas (PINEAU, 1998; DELORY-MOMBERGER, 2000; PADILLA, 2001); la investigación educativa (ABELS-EBER, 2000), la educación e intervención intercultural y su análisis discursivo (LERAY y BOUCHARD, 2000); etc.

Plantear la historia de vida, en cualquiera de las modalidades que se utilice como instrumento y método de diagnóstico e intervención, requiere plantear este enfoque desde la perspectiva de complejidad que conlleva, que se aproxima a una concepción de la ciencia, de la formación, de la evaluación y de la intervención, compleja, dinámica y viva en donde el desorden forma parte del orden y de la evolución y son características propias de la vida, como exponen ciertos autores (CAPRA, 1998; FORTIN, 2000, etc.).

> «Desde principios de siglo se conoce que el patrón de organización de un sistema vivo es siempre un patrón de red. No obstante, sabemos también que no todas las redes son sistemas vivos. Según Maturana y Valera, la característica fundamental de una red viviente es que se está produciendo a sí misma continuamente. Por tanto, "el ser y el hacer (de los sistemas vivos) son inseparables y éste es su modo específico de organización". La autopoiesis, el "hacerse a sí mismo", es un patrón de red en el que la función de cada componente es participar en la producción o transformación de otros componentes de la red, de tal modo que ésta se hace a sí misma continuamente. Es producida por sus componentes y, a su vez, los produce» (CAPRA, 1998: 175).

Este sentido de patrón de un sistema vivo en red que constantemente se está transformando y produciendo a sí mismo de manera dinámica, en donde el orden y el desorden forman parte de este proceso de la dinámica viva, sitúa a la vida y a la historia de vida en la perspectiva de la *ciencia de la complejidad*. Esta perspectiva representa una inversión de las visiones científicas tradicionales. Así, en la perspectiva clásica, se asocia orden con equilibrio, mientras que el desorden se identifica con situaciones de no-equilibrio.

Para CAPRA (1998), la nueva ciencia de la complejidad, que se inspira en la trama de la vida, el no-equilibrio es entendido como fuente de orden.

> «En los sistemas vivos, el orden emergente de no-equilibrio resulta mucho más evidente, manifestándose en la riqueza, diversidad y belleza del mundo que nos rodea. A través del mundo viviente, el caos es transformado en orden» (CAPRA, 1998: 202).

Este concepto de la ciencia de la complejidad, en donde el no-equilibrio es una fuente de orden nacida de la dinámica propia del orden-desorden, se recoge cuando se alude a creación y dinámica de la propia vida, física, biológica y humana. El orden y el desorden son inseparables para concebir la innovación, el progreso y la evolución. Es esta difícil confrontación la que tenemos que admitir para ser capaces de reconocer la complejidad de todo lo que es físico, biológico y humano (FORTIN, 2000).

Aludir a la historia de vida como método o instrumento de diagnóstico e intervención es situarse en las perspectivas de la concepción dinámica y compleja del ser humano, así como de un concepto de la ciencia global, en red, sistémica, dinámica, transformadora y evolutiva, que atraviesa las diversas dimensiones existenciales, sociales y ecológicas de las personas de manera dinámica e interactiva. En la medida en que las personas empiezan a escribir o narrar sus vidas, se produce un efecto de remover dinámico, que crea un desorden que afecta tanto a las dimensiones emocionales, como a las intelectuales y racionales, y desorganiza el orden estructural y sistémico en el que éstas están conformadas y se ubican.

Cuando una persona se siente incómoda con su vida, o con algunos aspectos de la misma, en los que siente una desarmonía, un cuestionamiento o un malestar que le crea problemas importantes, se preocupa por encontrar las causas o razones que le crean esa situación de malestar. Utilizar la historia de vida como un instrumento para detectar las causas de esa situación significa que tenemos que tener conciencia de la complejidad de este instrumento-método, que no es lineal, sino dinámico, complejo y global. Un método que, en la medida que interviene para narrar, describir situaciones o expresar vivencias, no sólo nos está aportando informaciones o retazos de información, sino que está produciendo una toma de conciencia en el sujeto que, a la vez que lo desestructura, le da información de su estructura, lo desorganiza, lo dinamiza, le permite tomar conciencia de quién es, y le facilita recursos para reestructurarse de otro modo.

Al utilizar la historia de vida es necesario no sólo conocer la naturaleza del método, sino tener conciencia de que esta intervención no es una cosa puntual, sino que afecta a la proyección del estilo o modo de vida que el propio sujeto tiene de sí mismo o desea desarrollar.

Según DELORY-MOMBERGER (2000), la formación debe permitir a la persona que hace su historia de vida descifrar su existencia, comprenderla y darle sentido. El sentido de la historia individual se elabora a partir de las significaciones que el sujeto atribuye a los acontecimientos de su vida. Esta forma de proceder se inscribe en una perspectiva emancipadora que tiene por objetivo permitir al sujeto situarse y estructurarse en relación a sí mismo y a los otros en el juego de las interacciones. Por otra parte, implica:

> «Hacer la apuesta de que el sujeto va a poder en lo sucesivo actualizar y desplegar sus potencialidades en un proyecto de acción más pertinente, que él podrá situarse de manera crítica en su entorno, convertirse en un actor social» (DELORY-MOMBERGER, 2000: 250).

Desarrollar un diagnóstico a través de la historia de vida exige tener conciencia de que es un proceso en el que se da simultáneamente la recogida de información con la intervención; un proceso que lleva al cuestionamiento, la interpelación, la clarificación, la toma de conciencia del problema, el propósito del cambio y la planificación global del proyecto de vida. Esto proporciona al sujeto una liberación de las estructuras coercitivas que le impiden desarrollarse, crecer y convertirse en una persona más autónoma. Llevar a cabo ese proceso es difícil, y abordar la escritura puede producir dolor pero, al mismo

tiempo, ayuda a descargar el peso que produce una situación, y a aligerarlo, como se refleja en el párrafo siguiente:

> «S., 11 años, escribe mucho en su cuaderno que le es muy personal y muy querido; "yo escribo todas mis miserias, y después, algunas veces, hay algunas cosas que yo escribía, que yo se las hacía leer a la asistente social y a la cuidadora, después arrancaba la página y la quemaba; le decía: ya está, ahora ya sabe lo que me pasa. Pero más tarde yo me digo, que no merece la pena quemarlo, porque es duro lo que yo he vivido, pero no solamente estoy yo, contaros todo eso me demuestra que no está mal lo que escribo, y esto me produce bienestar, porque es necesario que eso salga, esto me hace bien, es pesado llevarlo sola, y al mismo tiempo lo guardo para mí, esto me gusta, me gusta mucho mi cuaderno, es mi pequeña historia"» (ABELS-EBER, 2000:73).

Cuando la historia de vida se desarrolla en presencia de otro u otros adquiere una dimensión particular de *socialización*. Es esencial que las personas se sientan escuchadas comprendidas y valoradas. En el caso de que el proceso se lleve con un especialista, éste no sólo debe comprender la lógica de ese proceso y haber pasado por la reconstrucción de su historia de vida, sino que además es esencial que refleje su *autenticidad*.

> «El acompañante debe ser "congruente", para retomar el término de Rogers: capaz de ser la persona que es, ser él mismo en la animación con sus entusiasmos, sus rechazos, sus dificultades para ciertos temas, tesis y teorías que presenta y no la encarnación abstracta de una experiencia escolar y social. Debe autorizarse a LLORAR, REIR, AMAR, COMPRENDER, (...). Y debe haber vivido él mismo los procesos que propone a los acompañados. El aprendizaje auténtico supone también que el acompañante reconozca al acompañado, en una palabra que lo acepte tal como es. Lo que supone tener confianza en su capacidad para gestionar este proceso» (LIETARD, 1998: 117).

Cuando el proceso se desarrolla en grupo es muy importante establecer *unas normas de funcionamiento y unas normas éticas*. Muchas de estas normas también deben respetarse cuando el proceso se desarrolla entre el otro y el acompañante. La primera de ellas es el *concepto de comunicación*:

> «Comunicar es poner en común (...), ya sean nuestras diferencias o nuestras semejanzas. Expresarse no es comunicarse, es un ir simple, y la comunicación es un ir y venir que comporta unas etapas sucesivas:
>
> Yo me expreso.
>
> Yo recibo la confirmación de que mi lenguaje ha sido entendido.
>
> Yo escucho la expresión del otro.
>
> Yo le confirmo que lo he entendido.
>
> Estas secuencias comportan numerosas trampas. El otro no reacciona por lo que digo, sino por lo que entiende, según la resonancia que mi palabra tiene en él» (SALOMÉ y GALLAND, 1990:15).

Al intentar decir las cosas podemos decir, o no decir; hablar sobre nosotros mismos o sobre el otro, el mundo, la vida...; hablar de hechos, de sensaciones, sentimientos, etc. Según SALOMÉ y GALLAND (1990:11-20), la expresión verbal se puede situar en cinco niveles diferentes:

— El *nivel de los hechos*, que es el registro anecdótico que permite decir qué ha pasado y cómo.

— El *nivel de las sensaciones y sentimientos*, que es aquel de la vivencia y del sentimiento que acompaña a una situación o a un encuentro. Afecta a la zona sensible de las reacciones afectivas y dependerá de la disponibilidad de la persona a

exponer su vivencia íntima y a la manera en que el otro pueda entenderle y estimularle.

— El *nivel del pensamiento y de las ideas*, que es el de las generalizaciones, el de la evasión de la palabra hacia unos horizontes sin límites.

— El *nivel de la resonancia*, que se desarrolla como si cada acontecimiento del presente tuviera un eco en un pasado próximo o lejano.

— El *nivel de lo imaginario*, que es la fuente privilegiada de la creatividad en el diálogo.

Una comunicación plena es un compartir donde tengo la posibilidad de decirme, de ser escuchado y de escuchar al otro en estos cinco niveles, en los diversos momentos de una comunicación y de una relación. Las comunicaciones profundas evolucionan libremente de una dimensión a otra, y dejan circular la palabra en cada una de estas dimensiones.

Otro aspecto que nos parece importante considerar es lo que *no se dice*, ello significa que es importante aceptar los límites de la comunicación; es una especie de autorregulación de la comunicación que permite a cada uno salvaguardar ciertas zonas de intimidad.

> «Estar solo en presencia del otro. Sutil comunicación entre dos seres que se escapa al interior de ellos mismos, mientras que una fuerza magnética resulta de esas dos corrientes paralelas y distintas» (SALOMÉ y GALLAND, 1990:23).

La *escucha* es otro de los elementos a tener en cuenta. Escuchar sin apoderarse de la palabra requiere una disponibilidad y una descentración. Escuchar es acoger al que se expresa sin emitir juicios de valor, intentando comprender su mundo interior en su sistema de referencia. Al mismo tiempo, cuando el acompañante retoma o resume lo que el otro ha dicho, o lo que él ha entendido, pone en práctica la escucha activa, que es lo que permite al otro ir más lejos en su expresión, avanzar en su entendimiento de sí mismo (SALOMÉ y GALLAND, 1990).

Es importante *entender y comprender al otro en sus distintos lenguajes*; es decir, ir más allá de la escucha para captar lo esencial y poder distinguir en qué registro está hablando en cada momento y qué quiere comunicar con él. Esto requiere estar extremadamente atentos y descentrarnos de nosotros mismos, sabiendo distinguir lo esencial del mensaje y facilitarle al otro *parirse a sí mismo* a través de esta comunicación. Según SANTAMARINA y MARINAS (1994), el acompañante actúa como una comadrona que ayuda a parir no sólo un relato de la memoria o la experiencia, sino también ayuda a parir una representación del sujeto a partir de sí mismo.

Se consideran esenciales los *desfases y malentendidos* que suelen darse en la comunicación cuando se encabalgan los distintos niveles: afectivo, lógico y normativo. La lógica racional habla un lenguaje distinto al de la lógica afectiva, y a veces son irreconciliables. Es decir, sentimientos, pensamientos racionales y sistemas de valores

son registros que existen en todos y cada uno de nosotros; registros que a veces se cruzan en la comunicación produciendo malos entendidos.

Otro aspecto importante a considerar es el de la *comunicación consigo mismo*, ya quecomunicarse con el otro pasa por mejorar la comunicación consigo mismo. Cualquier proceso de cambio que pretenda mejorar la comunicación con el otro pasa primero por el cuestionamiento y reconocimiento de los propios sentimientos y vivencias. Estar lo más cerca posible de los propios sentimientos reales, estar a la escucha de las propias emociones exige mucha atención y vigilancia.

> «Reconocer mis sentimientos reales es reencontrar un estado de congruencia que va a confirmarme mi fiabilidad. Es también darme los medios de ser más coherente en mis conductas, en mi posicionamiento al frente del otro» (SALOMÉ y GALLAND, 1990:35).

En la dinámica de comunicación que se establezca, bien sea en grupo o entre el acompañante y el acompañado en el proceso de diagnóstico-intervención con la historia de vida, es importante respetar todas estas normas de comunicación, enviando reflexiones al otro, haciendo nosotros de espejo. Reflexiones e interrogantes que le lleven más lejos en su clarificación y reformulación, así como en su reestructuración. La puesta en común de diversas historias de vida hace emerger aspectos comunes y diferentes en todas y cada una de ellas, que nos sitúan en lo sociológico general y en lo específico existencial.

Otro aspecto que consideramos importante es el *posicionamiento ético de respeto al otro y respeto al secreto de la información*, clarificando lo íntimo que se considera de orden público y aquello que quiere conservarse en el orden de lo privado, respetando las decisiones individuales y colectivas a ese respecto y evitando las posibles repercusiones negativas que puedan surgir.

Por último, consideramos que el *análisis* de la diversa, variada y compleja información que va saliendo en el proceso de la historia de vida, debe realizarse teniendo en cuenta la complejidad global e interactiva de la persona como ente vivo, orgánico y sistémico, que pueda reflejar esta dinamicidad y clarificación. Es importante clarificar y analizar los aspectos temáticos, lingüísticos, afectivos, filosóficos, actitudinales, etc., que nos permitan reorganizar nuestro sistema mental.

> «Pienso que el problema crucial es el del principio organizador del conocimiento, y lo que hoy es vital, no es sólo aprender, tampoco reaprender ni desaprender, sino reorganizar nuestro sistema mental para reaprender a aprender» (FORTÍN, 2000:17).

En el análisis del relato producido por el sujeto o sujetos, es importante que sea releído de manera conjunta por éste/os y el acompañante del proceso. De esta actividad surgirán muchas precisiones sobre el significado del discurso del sujeto, así como una nueva ocasión de ampliar los detalles poco claros del relato, o cuestiones que pueden ser importantes y que no fueron mencionadas. En este sentido, como señala GONZÁLEZ MONTEAGUDO (1996), en todo documento biográfico hay que diferenciar siempre el contenido manifiesto del contenido latente.

Estos son pues algunos de los presupuestos que requiere la utilización de la historia de vida como instrumento y método de diagnóstico e intervención educativa.

3.3. Operativización de la historia de vida. Análisis de un ejemplo: *Trozos de mi vida*

Las formas estratégicas en las que se operativizan las historias de vida pueden ser muchas y variadas. Nosotros nos limitaremos aquí a plantear el proceso seguido por 14 de las 15 mujeres mayores que han desarrollado su historia de vida y que se recogen en el libro, elaborado por ellas, titulado *Trozos de Mi Vida* (CENTRO DE ADULTOS «JUAN XXIII», 1999:264-274).

Estas mujeres cursan el programa de Educación de Personas Adultas, por lo que la realización de su historia de vida y su publicación supone un reto importante para todas ellas: tienen que aprender a dominar la escritura, al mismo tiempo que la complejidad de la historia de vida y el manejo del ordenador para hacer su libro. Sin embargo, el proceso fundamental que requiere el desarrollar una historia de vida es común a la esencia misma de este instrumento llevado a cabo con personas que manejen las herramientas de la escritura y la informática. Por eso nos parece importante plasmarlo aquí. A continuación, iremos analizando algunas de las características más significativas del proceso desarrollado.

En cuanto al grupo que utiliza la historia de vida, éste está constituido por 15 alumnas que desarrollan 14 historias de vida. Todas ellas personas mayores que se encuadran dentro del nivel de Formación Inicial de Base. Este nivel se encuadra, a su vez, en el ciclo de Neolectores. Todas sus componentes cursan este nivel académico, exceptuando dos personas que ya tienen el título y que participan a nivel simbólico. La experiencia se desarrolla en el Centro de Adultos «Juan XXIII» de Sevilla.

Los *objetivos generales* que se persiguen con la utilización del instrumento son los siguientes:

— Desarrollar en las personas que componen el grupo una conciencia de que no son sólo una parte de la historia, sino que ellas son y hacen la historia, y que gracias a ellas ésta es posible.

— Escribir y publicar un libro con las historias de vida de cada una de las componentes del grupo.

— Formar un grupo de trabajo en el que la base fundamental sea el apoyo entre las componentes.

— Usar las nuevas tecnologías (ordenadores) para escribir su historia.

— Desarrollar una conciencia histórica colectiva que les permita comprender problemas actuales del mundo que vivimos y quiénes son sus protagonistas.

Estos objetivos generales del grupo se tenían que concretar en el trabajo a desarrollar en cada una de las áreas de estudio.

En cuanto al *proceso metodológico* seguido en la confección de los textos personales, lo primero que hay que resaltar es que el proceso duró todo un curso escolar. En las distintas sesiones, la tarea central ha sido la de redactar la historia personal de cada participante/alumna.

Dicha tarea fue estructurada cronológicamente, para facilitar su desarrollo, en un conjunto de apartados. A la hora de plantear el trabajo dentro de cada uno de estos apartados, se llevaban preparadas algunas preguntas para incitar al diálogo y unos elementos fundamentales que debían reflejarse a la hora de escribir sobre cada uno de ellos. Los apartados que han estructurado el proceso son los siguientes:

— *Nacimiento*. Año, lugar y circunstancias. Para suscitar la narración, se utilizaron textos generadores como *El hereje* de Miguel Delibes, o *La familia de Pascual Duarte* de Camilo José Cela.

— *Presentación de los padres.* Se debería convertir, junto con el apartado anterior, en la presentación de su familia y de las circunstancias por las que estaban pasando en el momento de su nacimiento. No se incluye como apartado la presentación o alusión a sus hermanos, sin embargo, para algunas personas es muy significativo e imprescindible hablar de ellos. Aquí también se utilizan textos generadores como el mencionado de *La familia de Pascual Duarte* y *Memorias del Sur* de Rafael Gómez Pérez.

— *Bautizo.*Fecha, iglesia en la que fueron bautizadas, padrinos y su parentesco, circunstancias, costumbres de la época con respecto a este acontecimiento, celebración...

— *Juegos.* Este apartado es muy importante ya que muestra el ambiente en el que se desenvuelven en la época infantil. Se tocan las relaciones con sus semejantes, lugares donde se encuentran, estilos de juego, materiales que emplean, de dónde sacan los materiales, juegos colectivos/individuales, canciones, relaciones con los mayores...

— *Colegio.*Tiempo de estancia en el colegio, nombre del colegio, tipo (público/privado), forma de pago en su caso, relación con los maestros, relación con los compañeros, recuerdos que les deja la experiencia escolar, material escolar, forma de conseguir el material escolar, actividades que se realizaban en el tiempo escolar...[31]

— *Primera Comunión*. Preparación, iglesia en la que se celebra, participación en la celebración, vestimenta (componentes, procedencia, uso, dinero que costó...),

celebración, costumbres de ese momento, etc.

— *Trabajo.*Se les motiva a describir su primer trabajo (dónde estaba situado, cómo accedieron, qué actividad desempeñaban, forma de desplazamiento, horario, sueldo, quién lo cobraba, etc.). Asimismo, se indaga sobre los siguientes trabajos (por qué se va del primero, en qué trabajaba, ventaja con respecto al anterior...).

— *Pretendientes, novios.*Cómo conocieron a su primera pareja, quién era, dónde trabajaba, por dónde salían, qué hacían, cómo eran las relaciones con los padres, etc.

— *Boda.* Tiempo de noviazgo, forma de decidirse a casarse, iglesia donde tiene lugar, celebración, lugar donde pasan la noche de boda, viaje de novios, casa donde van a vivir, condiciones de la vivienda, etc.

— *Trabajo del cónyuge.* Lugar, sueldo, profesión, condiciones de trabajo, horario, relación sueldo/necesidades familiares, etc.

— *Hijos.* Cuándo quedan embarazadas, tiempo entre el embarazo y la boda, condiciones de vida del momento, parto, condiciones del parto, costumbre de esta época, descripción de los hijos. Experiencias de otros hijos, relación con respecto al primero, etc.

Aunque todos los apartados estaban delimitados, el trabajo no se realizó de forma lineal, sino que para poder ayudarles a producir un texto a cada persona en cada uno de los apartados, era necesario generar primero un diálogo sobre el tema para motivarlas, además de usar otros recursos como textos de apoyo, canciones, etc. A partir de la lectura del texto motivador, las alumnas escribían sus vivencias sobre la temática correspondiente en su cuaderno. El trabajo se continuaba en casa y era apoyado por los coordinadores para que lo matizaran.

Una vez confeccionados los manuscritos, se transcribían en el ordenador. Aquí los responsables del grupo las iniciaban en el manejo del mismo y les corregían las faltas. Una vez corregidos e impresos, los textos eran devueltos a las alumnas. En una nueva sesión de grupo, eran leídos en voz alta, con la finalidad de que las alumnas se reconocieran en ellos, los matizaran, ampliaran y remodelaran.

Una vez realizados todos los textos completos, y según un guión previamente establecido, se procedió a realizar un análisis histórico de todos y cada uno de ellos, según los diversos aspectos tratados: la época, las circunstancias, los lugares, las costumbres, etc. Este análisis hace emerger los aspectos generales que han vivido como grupo, como parte de una generación histórica y una clase social, etc. Esto les lleva a una toma de conciencia histórica, a sentirse sujetos de la Historia y a aprender a relativizar ciertas situaciones. Este proceso ha sido puesto de manifiesto por otros autores como DE MIGUEL (1996), quien sugiere que las historias de vida son como *ventanas*, ya que gracias a ellas se pueden atisbar las normas, los valores y el contexto que prevalecían en un momento histórico concreto. Las mujeres que han vivido este proceso descubren

también sus aspectos y vivencias específicos y singulares, lo que les lleva a darse cuenta de su individualidad.

De esta segunda lectura de las historias de vida, surge la necesidad de profundizar en el análisis histórico y de ilustrar los textos con documentos (que buscan en la Hemeroteca), fotos, etc. Asimismo, esto conlleva otras tareas como elegir cómo y dónde poner estos materiales en el texto, qué estructura darle al libro, etc.

Como broche final, debaten sobre todos estos aspectos, presentan el libro a sus compañeras del Centro y hacen una evaluación del trabajo: del proceso seguido, las resistencias y bloqueos que se han manifestado, los momentos difíciles que se ha atravesado, los aprendizajes que se han adquirido, el esfuerzo y la implicación desarrollados, las satisfacciones, logros, tomas de conciencia, etc.

Es importante destacar que el desarrollo de estas historias de vida ha conllevado distintas dificultades y resistencias, así como diferentes niveles de implicación por la temática. Algunas mujeres manifiestan cierta resistencia a escribir algo de carácter personal, al esfuerzo de revivir y escribir las vivencias propias, y al nivel de profundización, expresión y matización de las mismas. Otras hacen referencia al nivel de interés y de implicación generado por la temática. Asimismo, se aprecian diferentes niveles de trabajo dentro del grupo; mientras unas avanzan rápidas, otras se resisten, incidiendo esto en la dinámica general del grupo, que se ha visto obligado a crear apoyos individualizados y a respetar ritmos y saberes plurales.

También se han generado procesos de aprendizaje cooperativo y colaborativo, existiendo una dinámica de apoyo y comunicación dentro del grupo. Las mujeres con mayor nivel ayudaban y estimulaban al resto y, a nivel contextual, se han buscado fórmulas de poder utilizar los recursos del Centro y desarrollar este trabajo específico en armonía, con la planificación de las distintas tareas del mismo.

A partir de este ejemplo se ve claramente cómo el desarrollo de una historia de vida es esencialmente procesual y requiere un acompañamiento personal y grupal que ayude a estimular a las personas, a vencer las resistencias, a compartir la intimidad, respetarla, reconocerla y reformularla de otro modo, así como a situarla en lo general y común, y en lo específico y particular.

4. EL GRUPO DE DISCUSIÓN

El grupo de discusión constituye una técnica usada ampliamente en el campo de la investigación de mercados, como procedimiento para explorar los valores y opiniones que distintos públicos tienen de un determinado producto. Progresivamente se fue introduciendo en la investigación sanitaria como un complemento a los estudios tipo encuesta. Así, se usaba el grupo de discusión como un paso previo para identificar los temas a incluir en un cuestionario (WARD, BERTRAND y BROWN, 1991). Sin embargo, esta técnica de recogida de información fue consiguiendo una cada vez mayor aceptación, introduciéndose en la investigación social, en general, y en la investigación educativa, en particular. WITKIN y ALTSCHULD (1995) afirman que, en el contexto estadounidense, se usa también como medio de asesoramiento para que fiscales y abogados defensores seleccionen a los posibles miembros del jurado.

En general, es una técnica que puede considerarse cualitativa, dado el tipo de información que genera y produce. Así, más que explicar hechos, el interés reside en comprender los fenómenos sociales desde la propia perspectiva del actor, lo que implica el uso de una metodología con clara orientación fenomenológica. Mediante el grupo de discusión no se buscan explicaciones causales descontextualizadas, sino reflejar la experiencia social que construyen, organizan y perciben las personas.

En realidad son muchos los campos del saber en los que se utiliza el grupo, y generalmente, con dos finalidades diferenciadas: por un lado, como medio de producción de información, bajo una finalidad diagnóstica o investigadora y, por otro, como instrumento de intervención, transformación y terapia (los grupos terapéuticos habitualmente usados en la psicología constituyen quizá el ejemplo más claro de esta segunda aplicación). La técnica de grupo de discusión se sitúa, no obstante, en la primera finalidad; es decir, constituye una estrategia de generación, producción y captación de información sobre una realidad.

Una de las primeras cuestiones que debemos abordar en relación con el grupo de discusión es la pluralidad de términos diferentes que se utilizan para su denominación y la dificultad de discernir, en ocasiones, si se habla o no del mismo procedimiento. Así, en la bibliografía sobre el tema encontramos en ocasiones expresiones «afines» a la de *grupo de discusión*: entrevistas de grupo, *focus group*, grupo focal, entrevista de grupo focal, etc.

El término en inglés que se utiliza para referir a esta técnica en la que un grupo de personas se reúnen para intercambiar sus opiniones sobre un tema es el de *focus group*,o también el de *focus group interview*. De ahí deriva el hecho de que en castellano encontremos en ocasiones a este procedimiento denominado como *grupo focal* o *entrevista de grupo focal*. Expresiones que podemos considerar, en términos generales, como sinónimas de grupo de discusión.

Al mismo tiempo, puesto que la técnica básicamente coincide, en la forma, con el proceso que se suele desarrollar en las entrevistas abiertas, la expresión *entrevista grupal* es también muy utilizada para referir al grupo de discusión. Si bien una entrevista grupal puede tomar varios formatos y no necesariamente el de un grupo de discusión, se puede destacar que el grupo de discusión tiene lugar bajo la forma de una entrevista abierta y no directiva a un grupo de sujetos.

No obstante, existe otra técnica de recogida de información basada en el grupo que sí presenta diferencias notorias respecto al grupo de discusión. Nos estamos refiriendo a la *técnica de grupo nominal* (DELBECQ, VAN DE VEN y GUSTAFSON, 1984). En ella, el calificativo «nominal» precisamente se usa para denotar que el grupo es tal «grupo» sólo a efectos nominales, de tal forma, que la generación de información se produce siempre de manera individual, y no como consecuencia de un debate de grupo.

Para empezar adentrándonos en el grupo de discusión, podemos caracterizarlo básicamente como una técnica no directiva a través de la cual se pretende producir, de forma controlada, un discurso de un grupo de sujetos (KRUEGER, 1991). Durante un tiempo limitado (generalmente, entre una y dos horas), los sujetos del grupo interactúan entre sí, debatiendo sobre el tema que les ha propuesto el moderador o persona que conduce al grupo.

En el grupo de discusión, el objetivo es dejar a las personas sugerir dimensiones y relaciones acerca de un problema, dimensiones y relaciones que difícilmente una sola persona puede establecer y pensar.

Para ORTI (1986), en la discusión grupal, el grupo es un marco para captar las representaciones ideológicas, valores, formaciones imaginarias y afectivas, etc. dominantes en un determinado estrato, clase o sociedad global. A través de la discusión grupal se puede tomar contacto con la realidad, o incluso hacer una representación teatral de la misma en condiciones más o menos controladas, en las que los miembros del grupo colaboran en la definición de sus propios papeles orientados por un director o coordinador del grupo.

CANALES y PEINADO (1994) sostienen que el grupo de discusión combina tres elementos de modo propio para producir una situación discursiva adecuada a la investigación o al diagnóstico:

a) El grupo de discusión *no es tal* ni antes ni después de la discusión. Su existencia se reduce a la situación discursiva. Esto le confiere un carácter artificial, en tanto que el grupo se constituye como tal a instancias del investigador o profesional del diagnóstico. En este sentido, el que no existan relaciones previamente construidas es fundamental para evitar interferencias en la producción del discurso.

b) El grupo de discusión *realiza una tarea*. Su dinámica simula la de un equipo de trabajo, ya que se orienta a producir algo para un objetivo determinado. Según

CANALES y PEINADO (1994), el grupo se halla constituido por la tensión entre dos polos: el trabajo (razón de su existencia) y el placer del habla.

c) El grupo de discusión *instaura un espacio de «opinión grupal»*. Se instituye como la autoridad que verifica las opiniones pertinentes, adecuadas o válidas. En él, los participantes hacen uso de un derecho al habla al emitir sus opiniones, reguladas en el intercambio grupal.

Una característica propia del grupo de discusión es que los que lo conforman se desconocen entre sí. Además, en su composición ha de procurarse que las personas seleccionadas reúnan características homogéneas en relación al tema objeto de diagnóstico o investigación. Gracias a este desconocimiento y, al mismo tiempo, similitud entre los interlocutores, se pretende que los sujetos que conforman el grupo se encuentren en un medio abierto y relajado, en el que sientan libertad para expresar sus ideas y opiniones ante el tema.

Si los objetivos del diagnóstico al que se aplica la técnica implican trabajar con poblaciones diversas, es preferible formar varios grupos de discusión en los que se encuentren representados todos los segmentos de población, pero nunca en un mismo grupo.

Existen estudios, por otra parte, que han comparado el grupo de discusión y la aplicación de cuestionarios (WARD, BERTRAND y BROWN, 1991). Los resultados de ambos procedimientos tienden a ser muy consistentes, por lo que el grupo de discusión puede constituirse en una alternativa viable en situaciones en las que, por cuestiones económicas, no es posible hacer una aplicación masiva de un cuestionario. Evidentemente, a través del grupo de discusión no es posible generalizar los resultados, pero esta técnica puede resultar muy útil para determinados objetivos (evaluar necesidades, valorar la eficacia de un programa, etc.).

En los próximos apartados profundizaremos en el proceso a seguir en la preparación y realización del grupo de discusión, así como en la figura del moderador, cuya experiencia y habilidad serán necesarias para la buena marcha del mismo.

4.1. El proceso de preparación y realización del grupo de discusión

Los grupos de discusión suelen formarse con un número reducido de participantes. Entre 8 y 10 personas es el margen aconsejable. Superar los 12-13 participantes no es recomendable, ya que en grupos tan numerosos, y al estar limitado el tiempo, es posible que no participen todas las personas o no se pueda profundizar en las respuestas.

Estos componentes del grupo son seleccionados de forma anónima, de acuerdo con las características necesarias para lograr la finalidad del estudio que se realiza. Frecuentemente, la selección de los miembros del grupo o grupos a formar requiere el

contacto con alguna institución u organización que facilita posibles candidatos para el estudio.

Por otra parte, como ya hemos indicado, es conveniente que los participantes no se conozcan previamente ni guarden relación entre sí. De esta forma, los sujetos irán elaborando con mayor libertad su discurso propio sobre el tópico a diagnosticar o investigar.

Los contactos iniciales para la preparación del grupo de discusión deben realizarse con una semana o más de antelación, a fin de dar tiempo suficiente a los participantes para que ajusten su agenda. Algunos autores (e.g. ORTI, 1986) recomiendan que los contactos iniciales con los participantes sean realizados por personas distintas al moderador. Es muy importante en este primer momento que no se produzca ningún tipo de «contaminación» que potencie que los sujetos asistan al grupo con una idea definida y sesgada. En general, el sujeto debe conocer lo menos posible del estudio que se está realizando, a fin de que no «prefabrique opiniones o posturas» (ORTI, 1986:182).

Por otra parte, a la hora de seleccionar a los sujetos hay que tener siempre presente que cuando se celebre el grupo de discusión probablemente no comparezcan todas las personas inicialmente contactadas. Por ello es recomendable ampliar el número de contactos iniciales (de 12 a 14), además de establecer un segundo contacto recordatorio el mismo día o el día antes.

En la selección de participantes, otra precaución que puede resultar de utilidad es que las personas seleccionadas no tengan experiencia en grupos de discusión. Esto permite una mayor espontaneidad en las reacciones y elimina el problema de participantes «profesionales», quienes podrían liderar y monopolizar la reunión. Utilizando el argot de la investigación (CANALES y PEINADO, 1994), nos interesan participantes «vírgenes», es decir, aquéllos que acuden por primera vez a un grupo de discusión.

Es evidente que otra condición importante para su realización se refiere a la elección de un lugar adecuado, en el que sea posible una disposición circular que facilite la participación de todos los sujetos, y en el que no existan ruidos de fondo que distorsionen o imposibiliten la conversación. Asimismo, ORTI (1986) subraya la importancia de elegir un lugar «simbólicamente» adecuado para las características del grupo, de manera que no se coarte la expresión (por ejemplo, sería inapropiado para una discusión de empleados de una empresa elegir un lugar de la misma). Por su parte, WITKIN y ALTSCHULD (1995) sugieren que el lugar utilizado para el grupo de discusión debe estar en la línea del estatus y las expectativas de los sujetos.

El proceso a seguir en la preparación de la discusión de grupo es muy similar al ya apuntado para las entrevistas abiertas, por lo que no lo repetiremos aquí. En el grupo de discusión, aunque se persigue minimizar la intervención del moderador, es necesario asegurarse de que los temas clave serán explorados, por lo que es conveniente utilizar un guión previamente elaborado.

Dicho guión, al igual que sucede en la entrevista abierta, no pretende constituirse en el eje en torno al cual se estructure la discusión, sino que consiste en una lista de áreas generales que deben cubrirse en cada grupo. En el cuadro 7 presentamos un ejemplo de posible guión que hemos utilizado en un trabajo previo (MORENO, ARANDA y MONTERO, 1993). Dicho estudio pretendía responder a la pregunta «¿por qué las mujeres que no buscan quedarse embarazadas no usan, o usan incorrectamente, los métodos anticonceptivos?», y fue elaborado a partir de una sesión preliminar con un grupo de mujeres que no pertenecían a ninguno de los grupos de riesgo abordados en el estudio. El objeto de este guión era disponer de un recurso para conducir la entrevista hacia los temas de nuestro interés, ya que en algunos momentos la conversación se podía desviar de los mismos.

Una vez realizados los preparativos y llegado el momento del encuentro del grupo, el moderador debe procurar no interactuar mucho con los participantes en estos momentos iniciales. Nuevamente se trata de no generar expectativas o dar información que sesgue su futura participación.

Tras sentarse todas las personas (en círculo, como ya hemos indicado), el moderador tendrá que hacer una introducción en la que se deben recoger varios aspectos:

— En primer lugar, se agradece la asistencia de las personas.

— Después se pasa a explicar brevemente los objetivos del estudio. Esta explicación debe ser veraz pero a la vez vaga e imprecisa, con el fin de reducir la inhibición de las personas y sesgar la información que éstas van a ofrecer en una dirección determinada.

— Se deben dar indicaciones sobre la dinámica a seguir. Estas indicaciones consisten en hacer saber a los participantes que son sus opiniones y valoraciones las que tienen mayor interés y que el moderador sólo actúa como conductor del grupo, nunca en calidad de experto o de participante en el mismo. Asimismo, se debe indicar que es conveniente que todos participen y guarden unas reglas básicas de funcionamiento (no interrumpir a los demás, no descalificar otras opiniones, etc.).

— Se debe hacer también referencia a la confidencialidad y privacidad de la información que aportan los participantes.

— En estos primeros momentos, es necesario solicitar la autorización para la grabación y justificar la necesidad de su uso.

Respecto a la grabación, su uso en el grupo de discusión suele ser necesario. Es la única forma de recoger íntegramente el discurso de los participantes. Si, por otra parte, el moderador tuviera que estar continuamente tomando nota, encontrará bastantes dificultades para moderar o dinamizar el grupo. Todo lo relativo a qué aspectos se deben tener en cuenta para la grabación, cómo es conveniente presentar esta a los participantes, etc. puede encontrarse en el capítulo de este libro dedicado a la observación, por lo que no lo mencionaremos aquí.

CUADRO 7

EJEMPLO DE GUIÓN UTILIZADO EN UN GRUPO DE DISCUSIÓN

- ¿Saben que son los métodos anticonceptivos?, ¿para qué sirven?, ¿cuál/es consideran mejor/es?
- ¿Cuál es el que usan? ¿por qué razón usan/no usan los métodos anticonceptivos?
- ¿Quién de los dos (la persona o su pareja) decidió utilizar ese método?
- ¿Qué ventajas tienen los métodos anticonceptivos?, ¿qué inconvenientes?
- Opinión de la pareja sobre los métodos anticonceptivos
- ¿A quién corresponde informar sobre los métodos anticonceptivos?, ¿quién/quiénes deben ser los responsables de impartir educación sexual, y concretamente información sobre los métodos anticonceptivos?
- ¿Quién les informó sobre los métodos anticonceptivos?, ¿consideran que están suficientemente informadas?
- ¿Conocen a alguien que, teniendo relaciones sexuales, no use los métodos anticonceptivos?, ¿qué le dirían a una mujer para convencerla de que es conveniente usar métodos anticonceptivos cuando tiene relaciones sexuales y no quiere quedarse embarazada?

(MORENO, ARANDA y MONTERO, 1993)

El moderador, no obstante, debe tomar algunas notas que le permitan complementar la información recogida en la grabación y, sobre todo, que le ayuden en su función de dinamizar el discurso. En ocasiones, es recomendable disponer de un co-animador que, además de liberar al moderador de la toma de notas, puede proporcionar un punto de vista diferente y complementario en relación con la discusión grupal.

Tras la presentación inicial, el moderador formula una pregunta muy general para iniciar la discusión, procurando que dicha pregunta no provoque una tendencia en los sujetos, es decir, que éstos no entiendan que tienen que responder en una dirección determinada o respecto a un contenido determinado. La formulación de la pregunta debe ser abierta y prestarse a la libre interpretación a la hora de responder. En el ejemplo de investigación sobre el uso de anticonceptivos, sería inadecuado empezar con una pregunta del tipo «¿qué métodos anticonceptivos usáis?», ya que ello supone considerar que todas las participantes usan métodos anticonceptivos. Una forma más adecuada de empezar podría ser: «Para empezar, ¿podemos hablar sobre los métodos anticonceptivos? Podéis mencionar todo lo que consideréis relevante en torno a ellos».

Se trata de que, a partir de la pregunta inicial, comience un debate sobre el tema propuesto en el que los participantes se expresen libremente y discutan sus opiniones y experiencias. Probablemente, la discusión grupal vaya abordando por sí misma todos los temas clave sobre los que se pretende generar información. Tan sólo en escasas

ocasiones será necesario sacar a relucir un tema o una idea que no ha sido explorada y debatida por el grupo. En ese caso, la forma de preguntar debe ser nuevamente abierta, de manera que no condicione las respuestas a dar, sino que la discusión del grupo fluya con libertad.

A lo largo del proceso de expresión, discusión y debate, el moderador ha de procurar intervenir lo menos posible. Sin embargo, en determinadas situaciones será necesaria esta intervención: cuando se producen largos silencios, cuando se genera una situación de excesivo conflicto, para animar a participar a los sujetos más callados, cuando se derive hacia otro tema, cuando alguien monopoliza la discusión, etc. En general, en estas intervenciones el moderador debe procurar no dar su opinión. Se trata más bien de que sea capaz de retomar las opiniones de los participantes para relanzar el tema o para animar a alguien a la participación.

La duración aproximada de un grupo de discusión puede variar aproximadamente entre una y dos horas. Menos tiempo indicaría que no se ha profundizado en el tema y más tiempo comenzaría a ser una cantidad excesiva por el cansancio de los participantes. En todo caso se recomienda que sea superior a 60 minutos, ya que es a partir de este momento cuando los grupos, una vez que han superado las resistencias iniciales, interactúan con mayor espontaneidad.

Cuando ha finalizado la discusión (tanto si se han desarrollado uno o varios grupos), se procederá a realizar la transcripción de la grabación para comenzar con las tareas de análisis. Aunque el material procedente de los grupos de discusión se puede someter al mismo proceso general a través del cual se analiza información cualitativa, BERTRAND, BROWN y WARD (1992) plantean otros posibles procedimientos a utilizar:

— Elaborar un inventario de los temas abordados. Tras la lectura del material se irían registrando todos los temas que van apareciendo en la discusión. Cuando éstos se repiten, se registran cruces en las correspondientes entradas. Este tipo de análisis es de gran utilidad para comparar las informaciones provenientes de varios grupos de discusión distintos.

— Recuerdo de los puntos tratados en el grupo de discusión por parte del moderador y posterior confirmación con las notas o la grabación. El moderador, por tanto, debe recordar, con el guión de la sesión por delante, las ideas principales aportadas en cada tema y corroborar su recuerdo con las notas ampliadas de la sesión. Este procedimiento tiene la ventaja de ser rápido y facilitar una primera toma de contacto con la información a analizar.

— Codificación al margen. Consiste en identificar, en el margen del texto de las notas o transcripción, unos códigos (numéricos o alfabéticos) que identifican los distintos temas abordados. Este último método evita la subjetividad inherente al anterior, además de consumir menos tiempo que la elaboración del inventario. Su principal dificultad estriba en la dificultad para organizar todos los posibles puntos abordados en cada tópico.

4.2. El rol del moderador del grupo de discusión

A pesar de que su nivel de intervención ha de ser mínimo, el moderador o dinamizador del grupo de discusión es una figura esencial. Las cualidades básicas que debe reunir pueden resumirse en dos: por un lado, una importante capacidad empática y, por otro, cierta experiencia previa en la conducción de grupos que garantice que no se dejará desbordar por la situación. Para conseguir ambos requisitos es necesario que la persona en cuestión haya participado antes en procesos grupales desempeñando en los mismos diversos roles.

Para facilitar el adecuado desarrollo de la discusión grupal, ORTI (1986) señala tres aspectos a tener en cuenta por parte del moderador: a) no formular juicios de valor sobre el tema; b) adaptar su lenguaje a las características del grupo; y c) mantener su autoridad moral.

Aunque no hay ninguna fórmula simple para asegurar el éxito, WITKIN y ALTSCHULD (1995) delimitan un conjunto de aspectos que el moderador debe tener en cuenta, a los cuales hemos añadido otros que consideramos relevantes:

— Comunicar comprensión y simpatía sin enjuiciar lo que dicen los participantes.

— Crear una atmósfera grupal agradable.

— Demostrar actitud de escucha.

— Ser capaz de llevar la conversación y no dejarse llevar por el grupo.

— Dejar que la conversación fluya.

— Comunicar un interés sincero en lo que los informantes están diciendo.

— Ser sensible al modo en que las propias palabras y gestos afectan a los participantes.

— Dejar claro, tanto implícita como explícitamente, que todas las opiniones son valiosas y que no se buscan opiniones o valoraciones en una dirección determinada.

— Estar alerta a los momentos en que el grupo se desvíe del tema objeto de discusión.

— Recordar los temas a abordar sin necesidad de estar demasiado pendiente de las notas.

— Resumir los progresos del grupo.

— Detectar cuándo la discusión está llegando a su conclusión o se agotan las energías del grupo.

Como ya se ha indicado, en la discusión de grupo hay ciertos momentos en los que se hace necesaria la intervención del moderador (largos silencios, clima agresivo, falta de

participación del algunos miembros del grupo, excesivo liderazgo de otros, etc.). En estas ocasiones, existen ciertas estrategias que el moderador ha de saber usar y aplicar en el momento apropiado. Algunas de estas estrategias son (PÉREZ CAMPANERO, 1991): la estrategia del *eco* (devolver preguntas o reflexiones que se acaban de formular), la estrategia del *rebote* (devolver a cada participante individual una pregunta o reflexión), la estrategia *reflectora* (en este caso, la pregunta o reflexión se devuelve a todo el grupo), etc.

Para completar este análisis de la figura del moderador del grupo de discusión, vamos a presentar aquí las sugerencias ofrecidas por MUCCHIELLI (1978) respecto a la actitud general que éste debe sostener en los procesos grupales:

a) Debe vigilar lo que sucede en el grupo. Esto es especialmente importante ya que *lo que pasa* en el grupo y *lo que se dice* no siempre son realidades coincidentes. Por ello, según este autor, el animador debe «comprender con dos registros». Este doble sentido es evidente en el silencio del grupo. La ausencia de la comunicación verbal implica algo a nivel de vivencia colectiva. El animador debe comprender el sentido de esta falta de palabras.

b) Debe permitir la espontaneidad y la participación de todos los implicados. Esto supone poner en práctica un conjunto de cualidades:

   — Demostrar soltura personal ante el grupo como una actitud que facilita la comunicación. Es fácil sentir miedo ante una situación grupal. A veces este miedo se manifiesta en la evitación de las situaciones de grupo, en la exageración de la soltura personal ante el grupo demostrando excesiva familiaridad o locuacidad, en la intensificación de las reacciones de autoridad, etc. Por el contrario, demostrar una soltura normal ante el grupo contribuye a denotar la capacidad de escucha y la fe en las propias capacidades del grupo para resolver una situación.

   — Demostrar una actitud que facilite el establecimiento de la relación y el «deshielo» del grupo.

c) Promover la no-directividad en el fondo y la directividad en la forma. No-directividad en el fondo significa que el animador no expresa opiniones personales, no emite juicios sobre las opiniones expresadas, no favorece una opinión respecto a otra... Por otra parte, la directividad en la forma hace referencia a la actitud vigilante que tiene que denotar el animador para que los participantes no se salgan del tema, aclarar lo que se dice, hacer síntesis en los momentos oportunos...

4.3. Consideraciones finales

Aunque nos hemos detenido en una única estrategia grupal para el diagnóstico, es necesario precisar que existen otras muchas técnicas que pueden contribuir a la generación de información significativa para el mismo. Entre ellas, destacamos los procedimientos de dinámica de grupo (*brainstorming*, Phillips 6/6, *role-playing*, mesa redonda, etc.), que se han utilizado habitualmente para fines didácticos más que propiamente evaluativos.

En muchas ocasiones los procesos diagnósticos adolecen de falta de imaginación y creatividad. Los profesionales del diagnóstico acudimos siempre a las mismas técnicas y no pensamos que pueden existir procedimientos alternativos que, precisamente por ser «diferentes» a lo habitual, generan una mayor motivación en los sujetos que participan en el diagnóstico. Si tenemos que evaluar las necesidades de orientación que presentan un grupo de estudiantes, por poner un ejemplo, acudimos por lo general al diseño de un cuestionario. No se nos ocurre que estrategias como las dinámicas grupales pueden servir a la misma función o aportar una información complementaria. Probablemente los alumnos afectados se sientan más motivados para participar en varias dinámicas grupales que para rellenar un cuestionario más.

Por otra parte, otro aspecto positivo que cabe destacar de las estrategias grupales —desde las más simples, como las mencionadas dinámicas grupales, hasta las más sofisticadas, como la técnica del grupo *delphi*— es que propician una participación más activa de los sujetos del diagnóstico. Esto conlleva que dichos sujetos se impliquen más activamente en el mismo. Así, cuando finalice el proceso diagnóstico y, como consecuencia del mismo, se propongan un conjunto de cambios o intervenciones a desarrollar, esta implicación de los sujetos será un elemento muy valioso para asegurar el éxito de la intervención, ya que los sujetos sentirán que son ellos quienes han producido la información, los verdaderos actores del proceso.

Por último, queremos hacer referencia a la cuestión de la ética en el uso y aplicación de las estrategias grupales. Estas normas éticas son necesarias siempre que se utiliza un instrumento o técnica de diagnóstico, sin embargo, en aquellos procedimientos en los que la participación de los sujetos es mayor, las precauciones a tener en cuenta también deben ser mayores. No es lo mismo rellenar un cuestionario —que, al fin y al cabo, no nos quita tanto tiempo— que participar en un grupo de discusión —donde, entre otras cosas, verbalizamos nuestras opiniones a la vista de un grupo de personas que no conocemos—. Algunas de las obligaciones éticas que el profesional del diagnóstico debe poner en práctica cuando usa estrategias grupales son: compartir sus resultados con el grupo, especialmente si éstos pueden beneficiar a sus miembros; asegurar y cumplir con la confidencialidad de la información; negociar, en caso necesario, el informe de investigación o diagnóstico, etc.

5. LAS REJILLAS DE CONSTRUCTOS PERSONALES

La peculiaridad que ofrecen las rejillas de constructos personales es que permiten que el sujeto genere su propia matriz de información. Así, partiendo de que cada persona valora e interpreta la realidad de forma distinta, se puede considerar que los factores que son significativos para un sujeto no tienen por qué serlo para otro. Por ello, en las rejillas de constructos personales, la persona trabaja siempre desde su marco interpretativo de la realidad: no sólo hace valoraciones de un conjunto de elementos, sino que también decide cuáles serán los elementos a ser valorados y los criterios utilizados para dicha valoración.

Esta técnica tiene su origen en la *Teoría de los Constructos Personales* de G. A. Kelly, quien la aplica al estudio de las relaciones interpersonales. Sin embargo, estos planteamientos iniciales se han extendido y, tanto la teoría como la técnica de este autor, tienen numerosas aplicaciones en la psicología y la pedagogía. Antes de introducirnos en la técnica en cuestión, consideramos necesario hacer una breve exposición de los fundamentos teóricos desde los que nace la técnica.

Según KELLY (1955), las personas construyen su mundo personal mediante su interpretación de los acontecimientos. De esta forma, un sujeto aprehende la realidad de acuerdo con sus sistemas de construcción, que dependen del tipo de supuestos en los que se basa y los valores que le asigna a las cosas.

La Teoría de los Constructos Personales se estructura en once corolarios que desarrollan el postulado fundamental según el cual los procesos de una persona se canalizan psicológicamente por la forma en que ésta anticipa los acontecimientos. Mediante el constructo personal, es decir, la captación simultánea de la similitud y diferencia entre acontecimientos, la persona comprende la realidad y anticipa el futuro (FEIXAS y VILLEGAS, 1990). Dicho de otra forma, el ser humano desarrolla constructos que son ejes de referencia o dimensiones de valoración que le proporcionan sendas de acción. Estos constructos son privativos de cada persona y tienen un carácter bipolar o dicotómico. Así, el sujeto es un científico que elabora hipótesis para interpretar y dar sentido al flujo de los acontecimientos en que se encuentra.

Aunque aquí vamos a centrarnos en la técnica de rejilla, en realidad la evaluación de los constructos personales puede realizarse mediante *entrevistas* (a partir de las cuales se confecciona la rejilla) y también mediante *textos y documentos personales de carácter autodescriptivo y autobiográfico*. Para profundizar en la evaluación de constructos personales mediante textos autobiográficos, recomendamos los trabajos de BOTELLA y FEIXAS (1990 y 1993) y FEIXAS y VILLEGAS (1990). Estos trabajos muestran cómo los textos autobiográficos pueden ser sometidos a análisis rigurosos que permiten evidenciar, de forma sintética y gráfica, las relaciones entre las dimensiones principales que el sujeto emplea para estructurar su mundo personal. Básicamente, se trata de

configurar el mapa esquemático y directo del sistema de construcción del sujeto, proporcionando una imagen global y sintética del mismo.

Centrándonos ya en la técnica de rejillas de constructos personales (RCP), y como hemos apuntado, KELLY (1955) la diseñó inicialmente con la finalidad de estudiar las relaciones interpersonales que construyen los sujetos, aunque con posterioridad la técnica ha sido utilizada para otros fines (otras aplicaciones de la psicología clínica, estudios de mercado, en la elección de ofertas formativas y profesionales, etc.). En general, la técnica permite explorar los sistemas de constructos personales o, como sugieren FRANSELLA y BANNISTER (1977:5), constituye una forma de «ponerse en la piel de otras personas, ver sus mundos tal y como ellas los perciben, comprender su situación, sus preocupaciones e intereses».

A través de una entrevista, la construcción y aplicación de la técnica RCP se desarrolla en un proceso que consta de cuatro pasos fundamentales (FERNÁNDEZ BALLESTEROS, 1992):

1. Elección de los elementos a utilizar (personas, eventos, objetos, etc.).

2. Establecimiento de los constructos relevantes para los elementos.

3. Representación de los datos.

4. Análisis de la información.

En la versión original de la técnica, los elementos estaban constituidos por personas relevantes para el sujeto, previamente elegidas de un listado amplio en el que figuran roles como la madre, el padre, la pareja, etc. Así, el sujeto elige cuáles son los elementos que va a considerar, tanto en lo que respecta al rol que desempeñan, como, dentro de un rol determinado, al nombre de la persona concreta que va a introducir en la rejilla.

Los constructos, por otra parte, se refieren a las dimensiones a partir de las cuales cada sujeto caracteriza, describe y valora los elementos considerados. Estos constructos son de carácter bipolar o dicotómico (como por ejemplo, divertido-aburrido, apoya-no apoya, etc.).

Dados los elementos y los constructos, la rejilla es, esencialmente, una matriz de datos; una tabla de doble entrada en la que los elementos constituyen las columnas y los constructos las filas. En el formato original propuesto por KELLY (1955:270), el sujeto debe decidir qué elementos reúnen un polo de cada constructo. El procedimiento es algo complejo, pues la cumplimentación de la rejilla se realiza al mismo tiempo que la explicitación de los constructos. Así, existen unos círculos en algunas casillas que indican qué elementos debe comparar el sujeto a fin de obtener los polos de semejanza (lo que es igual y comparten dos elementos del sistema) y de contraste (lo que es diferente al polo de semejanza y se encuentra en un tercer elemento del sistema). Posteriormente, el sujeto debe aplicar el polo de semejanza al resto de los elementos y

marcar las casillas según considere que a un elemento determinado se le puede aplicar el polo de semejanza.

Sin embargo, la técnica ha sido modificada por numerosos autores y puede hoy día adoptar diferentes formatos que permiten estudiar con más profundidad y precisión las relaciones entre los constructos que utiliza el sujeto. Algunos de los formatos posteriores (FRANSELLA y BANNISTER, 1977) son la *rejilla de ordenación jerárquica* (se ordenan los elementos de forma jerárquica en función del constructo), la *rejilla de puntuaciones* (se valoran los elementos según se acerquen a un polo u otro del constructo), la *rejilla de implicación* (en la que se aprecian variaciones entre constructos), etc. Asimismo, los elementos pueden presentarse también bajo otras fórmulas como fotografías, dibujos, etc.

Dado que el procedimiento de KELLY (1955) es bastante complejo y actualmente se utilizan más las variaciones sobre el mismo que el procedimiento original, en el siguiente apartado vamos a describir con más detalle el proceso de construcción de la rejilla de constructos personales utilizando concretamente la rejilla de puntuaciones.

5.1. Preparación, aplicación e interpretación de la técnica de rejilla de constructos personales

Como ya hemos indicado, la técnica de rejilla de constructos personales (RCP) se utiliza para conocer cuáles son los constructos que una persona emplea para describir, ordenar y dar sentido a la realidad. Aunque existen diferentes formas de aplicación de la técnica, ésta se emplea con los elementos y constructos que aporta cada sujeto individual. Esto hace que la rejilla, como instrumento de evaluación, se diferencie mucho de los tests tradicionales, los cuales se aplican de forma estandarizada y se orientan a medir los constructos, tal y como los concibe teóricamente el autor del test. La técnica de rejilla, por el contrario, se aplica de forma diferente en cada caso, según las finalidades del diagnóstico en que se utiliza. Por otra parte, son los constructos y teorías implícitas desarrollados por el sujeto los que se pretenden evaluar.

Aunque, en cierta forma, es empezar por el final, creemos que una primera visualización de un ejemplo de rejilla de constructos personales puede facilitar al lector comprender mejor el proceso que aquí vamos a explicar. Por ello, en el cuadro 8 se muestra una rejilla vacía en la que se aprecia cómo existen una serie de columnas y filas dentro de las cuales el sujeto tendrá que explicitar y organizar un conjunto de informaciones que, básicamente, se refieren a: los elementos, los constructos y las relaciones entre elementos y constructos.

CUADRO 8

EJEMPLO SIN CUMPLIMENTAR DE REJILLA DE CONSTRUCTOS PERSONALES

POLO IZQUIERDO DE LOS CONSTRUCTOS	ELEMENTOS										POLO DERECHO DE LOS CONSTRUCTOS

Como indicamos anteriormente, el proceso de construcción de la rejilla pasa por cuatro fases que son: la elección de los elementos, la elección de los constructos, la representación de los datos y el análisis de la información. Estas tareas se van realizando dentro del marco de una entrevista en la que se va desarrollando el proceso que dará lugar a la matriz de información. No obstante, existen ciertas aplicaciones de la técnica, como las que veremos en el apartado siguiente, que permiten una aplicación colectiva y la construcción de la rejilla en el marco de un proceso grupal. A continuación, haremos una descripción general de las distintas fases implicadas en la construcción de la rejilla.

El primer paso del proceso consiste en la selección y explicitación de los elementos. Según RIVAS y MARCO (1985), los elementos pueden ser diferentes cosas: pueden referirse a personas, a situaciones o eventos, a objetos físicos, preferencias, actitudes, etc.; «cualquier estímulo puede constituirse en elemento de una situación determinada, siempre y cuando sean susceptibles de conceptualización a través de constructos» (RIVAS y MARCO, 1985:88). Suelen ser también de gran interés los elementos del «sí mismo» que representan aspectos parciales o roles del sujeto como el «yo ideal», el «yo visto por los demás», el «yo como estudiante», etc. Según BOTELLA y FEIXAS (1998) y FEIXAS y CORNEJO (1996), es aconsejable dividir el elemento «yo» en varios aspectos posibles

como los ejemplos planteados. No obstante, los autores citados aconsejan no multiplicar en exceso este tipo de elementos, pues puede conllevar la exclusión de otros elementos significativos.

RIVAS y MARCOS (1985) clasifican los posibles tipos de elementos en las siguientes categorías: elementos de rol interpersonal, impersonales, de autoconcepto, de localizaciones, de objetos, de situaciones, ocupacionales, de estados internos, etc. No obstante esta diversidad en cuanto a la tipología de elementos que se pueden utilizar, KELLY (1955) impone dos condiciones o criterios que éstos deben cumplir:

— Por un lado, los elementos deben ser de tal naturaleza que les sean aplicables los constructos personales.

— Por otro, deben ser representativos del conjunto del que han sido extraídos y del dominio que se pretende evaluar.

Además de estos criterios, FEIXAS y CORNEJO (1996) aconsejan que los elementos sean: *homogéneos*, con el fin de que los constructos empleados puedan pertenecer a un mismo ámbito (por ejemplo, que todos los elementos sean miembros del grupo de relaciones cotidianas del sujeto y no haya alguien extraño al mismo); y *comprensibles* para el sujeto y adecuados a sus conocimientos.

A pesar de estas restricciones, el hecho de que existan tantas posibilidades a la hora de seleccionar los elementos hace que la técnica sea muy versátil y aplicable a distintos tipos de situaciones y problemas. En la formulación original de KELLY (1955) se sugería una lista de 24 roles[32], entre los cuales el sujeto escogía los de las personas más importantes en su vida. En las aplicaciones de esta técnica a los estudios de mercado, se escogen como elementos posibles productos (marcas de bebidas, tiendas de ropa, etc.). Si la aplicamos para facilitar la elección profesional, los elementos son profesiones o salidas profesionales; etc.

Por lo tanto, los elementos serán diferentes según el ámbito concreto en que se aplica la rejilla y las finalidades del proceso diagnóstico en las que esta técnica se usa. Esto también condicionará quién elige los elementos que se van a considerar en la rejilla. Así, en determinadas aplicaciones, por ejemplo en el campo de la psicología clínica o la investigación de mercados, es habitual que el evaluador seleccione los instrumentos a considerar. En otras aplicaciones, es el sujeto el que los selecciona, bien eligiendo todos los elementos, bien eligiendo la mayoría y sugiriendo el evaluador el resto.

En cuanto al número de elementos a considerar, éste dependerá de cada aplicación concreta. FEIXAS y CORNEJO (1996) aconsejan que, para revestir de mayor validez un diseño de rejilla, el mínimo de elementos se situaría en 10 y el de constructos en 8, siendo recomendable que, como mínimo, se obtenga un número superior a 100 al multiplicar los constructos por los elementos.

Supongamos que queremos utilizar la rejilla para explorar los constructos usados por una persona para describir sus relaciones interpersonales. Podemos tomar como

referencia la lista de KELLY (1955) o simplemente pedir al sujeto que los elija entre los roles más significativos de su entorno relacional. Si hay alguna persona que, según lo que conocemos previamente del sujeto, conviene ser incluida y el sujeto no la propone, el evaluador podrá proponerla. En el cuadro 9 aparece una rejilla que sirve para ilustrar este ejemplo (constructos implicados en las relaciones interpersonales de un hipotético sujeto X). Allí se aprecia que se han elegido personas significativas para este sujeto (su pareja, los miembros de su familia, un amigo...), además de considerarse el «yo actual» y el «yo ideal» y una persona *non grata*.

Una vez seleccionados los elementos, comenzará el proceso de selección y definición de los constructos personales. KELLY (1955) propuso diferentes formas para obtener constructos personales. Algunas de ellas son las siguientes:

— Procedimiento de *contexto mínimo*. El sujeto tiene que elegir tres de los elementos que ha considerado en la fase anterior. Ha de compararlos de forma que diga qué característica común poseen dos de ellos y los diferencia del tercero. La característica común se denomina constructo de semejanza y la del tercer elemento es el constructo de constaste.

— Procedimiento *secuencial*. Al igual que en el procedimiento de contexto mínimo se presentan tríadas variando cada vez sólo un elemento.

— Procedimiento de *autoidentificación*, semejante al anterior. En este caso el elemento «yo mismo» siempre forma parte de la tríada.

— Procedimiento de *contexto total*. En este caso se le presentan varias tarjetas con el nombre de las personas que desempeñan los roles elegidos por el sujeto en la fase anterior. Se trata ahora de que el sujeto forme grupos con las tarjetas y señale qué caracteriza a cada uno de los grupos formados.

BOTELLA y FEIXAS (1998) apuntan que este procedimiento —mediante tríadas— puede resultar complejo, por lo que proponen[33] utilizar parejas de elementos en lugar de tríadas. Así, presentada una pareja de los elementos seleccionados, se pregunta al sujeto por la característica que comparten, para posteriormente preguntarle cuál es, en su opinión, lo contrario de dicha característica. En caso de dificultad para discernir la similitud, se puede preguntar por la diferencia. Es necesario recordar que los constructos son bipolares y se trata de elicitar los dos polos del mismo.

En la presentación de las parejas de elementos, FEIXAS y CORNEJO (1996) aconsejan incluir cada elemento en al menos una de las parejas presentadas, así como incluir el elemento «yo» junto con el padre, la madre, la pareja, la persona *non grata* y algún hermano. En otro tipo de aplicaciones, el evaluador deberá decidir cuáles son las parejas de elementos que pueden sugerir dimensiones más significativas. Por otra parte, el procedimiento de elicitación de constructos a través de díadas o tríadas se repite hasta que llega un momento en el que al sujeto le resulta difícil elicitar nuevos constructos.

CUADRO 9

EJEMPLO DE REJILLA DE CONSTRUCTOS PERSONALES: LA REJILLA DE X

	JUAN (padre)	LUISA (madre)	MANUEL (hermano)	CARLOS (pareja)	MELI (amiga)	ELISA (compañera)	ANTONIA (*non grata*)	JOSÉ (admirado)	X	X ideal	
INCONSTANTE	3	2	2	1	2	4	4	5	4	5	SISTEMÁTICO
DECIDIDO	1	4	3	5	4	5	3	3	1	1	INDECISO
TOLERANTE	4	2	2	2	4	5	5	3	4	3	INTOLERANTE
HIPÓCRITA	4	4	5	3	3	4	1	4	5	5	SINCERO
REFLEXIVO	2	5	5	1	2	3	2	1	5	2	IMPULSIVO
DISFRUTA DE LA VIDA	2	5	1	4	1	2	5	2	4	1	NO DISFRUTA DE LA VIDA
AYUDA A LOS DEMÁS	5	1	1	4	3	3	5	4	4	3	NO AYUDA A OTROS
DEPENDIENTE	5	1	2	2	4	3	2	5	5	5	INDEPENDIENTE
MESURADO	4	3	4	2	3	5	1	2	5	4	RADICAL
COMPETENTE	2	2	1	1	1	2	1	1	2	1	INCOMPETENTE

1. Muy como el polo izquierdo.
2. Bastante como el polo izquierdo.
3. Punto medio.
4. Bastante como el polo derecho.
5. Muy como el polo derecho.

Sea cual sea el procedimiento utilizado para formular los constructos, es necesario tener en cuenta que se han de evitar ciertos constructos, repitiendo el proceso en caso de que esto sucediera, a fin de obtener constructos significativos. Los rasgos a evitar en los constructos, según FERNÁNDEZ BALLESTEROS (1992) son: a) que sean demasiado permeables (se pueden aplicar a un número excesivamente alto de elementos como «bueno-malo») o impermeables; b) que sean situacionales, es decir, que hagan referencia al lugar de origen del elemento (como, por ejemplo, un natalicio); c) que sean excesivamente superficiales; d) que sean excesivamente vagos; y e) que sean productos del título de rol.

En coherencia con el planteamiento teórico en el que se basa la RCP, los constructos han de ser elicitados por el sujeto a partir de la comparación de elementos. Sin embargo, el evaluador puede introducir un constructo si lo considera relevante en función de su conocimiento previo del sujeto (BOTELLA y FEIXAS, 1998).

En nuestro ejemplo del cuadro 9 podemos observar que los constructos aportados por X son «inconstante *versus* sistemático», «decidido *versus* indeciso», «tolerante *versus* intolerante», «hipócrita *versus* sincero», etc.

Seleccionados y formulados los elementos y los constructos por los que se van a valorar, se procede a la representación de la información a través de la técnica de rejilla. De esta forma se pueden representar y clasificar las relaciones entre constructos y elementos. Ya indicamos anteriormente que existen diferentes procedimientos para relacionar constructos y elementos. Aquí nos vamos a centrar en el tipo de rejilla que anteriormente denominamos como rejilla de puntuaciones.

La rejilla de puntuaciones es uno de los formatos que más se utiliza y que más fácil resulta de aplicar e interpretar. Este procedimiento consiste en dar a cada elemento una puntuación ordinal cuando se compara con el constructo. En este caso, hemos utilizado el procedimiento de escala tipo Lickert de 5 puntos propuesta por BOTELLA y FEIXAS (1998), sin embargo, se podría también haber elegido una escala de 3 o de más puntos.

En el ejemplo del cuadro 9 aparece la puntuación que el sujeto X va dando a cada elemento según al extremo de cada constructo que considera que se aproxima. Por ejemplo, a Juan (su padre) lo ha definido en el punto medio de «inconstante-sistemático», como muy decidido, bastante intolerante, bastante hipócrita, bastante reflexivo, etc. X, por otra parte, se ha autocaracterizado como bastante sistemático, muy decidido, bastante intolerante, muy sincero e impulsivo, etc.

El procedimiento se parece bastante en su forma a un diferencial semántico, donde el sujeto tiene que calificar a una persona, objeto o situación en función de pares de adjetivos. Sin embargo, no se debe olvidar que, a diferencia del diferencial semántico, en la rejilla de puntuaciones el sujeto es quien decide las dimensiones de evaluación, además de los elementos a que ésta se aplica.

La principal utilidad de la técnica es que ofrece la posibilidad de comparar todos los elementos (en el ejemplo 9, estos elementos están formados por personas que desempeñan roles significativos para el sujeto) con los constructos. Además, a través del análisis de las respuestas se puede obtener información acerca de cómo se relacionan los constructos entre sí dentro del sistema del sujeto.

Se trata ahora, por tanto, de realizar el análisis de la información vertida por el sujeto en la rejilla. Respecto a dicho análisis, lo primero que podemos decir es que, dada la gran versatilidad de la técnica (distintas formas de representar elementos, distintas formas de obtener constructos, distintas formas de relacionar constructos-elementos), los procedimientos analíticos que se pueden emplear son muchos y variados. Estos procedimientos pueden clasificarse, *grosso modo*, en dos tipos: por un lado, el análisis de contenido que puede aplicarse a fin de examinar los constructos en tanto que dimensiones psicológicas, y por otro, las herramientas estadísticas que van a facilitar el análisis de las relaciones entre constructos.

Fue KELLY (1955) quien utilizó de forma preferente el análisis de contenido. Su interés residía en conocer qué dimensiones construye el sujeto y cómo interpreta sus relaciones interpersonales a través de ellas. Por ello, sus análisis se enfocaban a conocer el número de constructos para estudiar la complejidad cognitiva del sujeto, el tipo y carácter de los constructos (si hacen referencia a dimensiones físicas, psicológicas, etc.), la relación entre los constructos que el sujeto se autoaplica y los que aplica a los demás, etc. En el libro de BOTELLA y FEIXAS (1998) se describen un buen número de posibles criterios a considerar y se detallan algunos consejos para su análisis e interpretación.

Asimismo, dado que la información recogida en la rejilla es también cuantitativa, es posible aplicarle estrategias de análisis cuantitativo. Algunos de los aspectos que se pueden analizar (BOTELLA y FEIXAS, 1998) son: relaciones entre constructos, relaciones entre elementos, los elementos en el espacio de los constructos y la construcción de la identidad del *self* o autoconcepto. A ello podemos sumar otras medidas propuestas por FRANSELLA y BANNISTER (1977) como las de diferenciación cognitiva (intensidad y complejidad), integración cognitiva y conflicto.

Aunque un primer factor que determinará el análisis a realizar es el tipo de puntuaciones que suministra la rejilla, que pueden ser dicotómicas (como la que utiliza KELLY, 1955), ordinales (los elementos aparecen ordenados en relación con el constructo) o de intervalo (en este sentido, las rejillas de puntuaciones como el ejemplo del cuadro 9 se consideran de intervalo). Así, se han propuesto y utilizado procedimientos analíticos como el análisis factorial no-paramétrico, el análisis de componentes principales y el análisis de cluster. Sin embargo, el análisis de correspondencias es el método que, según FEIXAS y CORNEJO (1996), permite una interpretación mucho más global y precisa, a la vez que más fiel a los datos originales. Actualmente existe un programa, denominado RECORD[34] (FEIXAS y CORNEJO, 1996), que realiza los cálculos necesarios para el análisis de correspondencias además de presentar algunas adaptaciones gráficas muy útiles para la interpretación de los resultados de las rejillas de constructos personales.

En cualquier caso, el análisis de la rejilla requiere, no sólo el dominio de las pruebas estadísticas necesarias, sino un profundo conocimiento de la técnica de la rejilla en sí misma, de la Teoría de los Constructos Personales en cuyo contexto se inscribe y, por supuesto, un conocimiento exhaustivo del ámbito de diagnóstico concreto al que la técnica se aplica. Por ello, en el apartado siguiente, presentamos un formato de rejilla de constructos personales más fácil de aplicar e interpretar para un profesional del diagnóstico que no esté familiarizado con la técnica.

5.2. Un ejemplo de rejilla de constructos personales aplicada a la decisión vocacional

Algunos ejemplos de rejillas construidos para fines concretos y, más o menos, sencillos de aplicar e interpretar, son los de ALONSO TAPIA (1997)[35] y ÁLVAREZ ROJO (1997)[36]. En el primero de ellos (ALONSO TAPIA, 1997) se plantea un ejemplo en el que el sujeto debe reflexionar sobre distintos tipos de estudio sobre los que puede decidirse; en este caso, se utilizan como constructos los criterios a través de los cuales el sujeto puede decidir entre dichos estudios (interés, duración, etc.). En el ejemplo aportado por ÁLVAREZ ROJO (1997) la rejilla se aplica como medio para facilitar la decisión entre un conjunto de profesiones y un conjunto de valores; ambos conjuntos (profesiones y valores) son aportados por los sujetos tras otras actividades anteriormente realizadas.

Aquí vamos a utilizar un ejemplo muy similar en contenido, aunque no en forma, al de ÁLVAREZ ROJO (1997). Este ejemplo ha sido desarrollado por MAU (1997) y ha sido utilizado ya por nosotros en otra publicación (PADILLA, 2001). Su selección se debe a que es una aplicación que combina la sencillez en cuanto a su aplicación (de hecho, está pensado para aplicaciones colectivas), con una cierta complejidad y profundidad en las puntuaciones que ofrece. A ello debemos unir el hecho de que está pensado para facilitar la toma de decisiones profesionales, lo que lo hace sumamente aplicable a las situaciones de diagnóstico educativo.

Como ya hemos planteado, esta rejilla pretende explorar las preferencias profesionales de los sujetos y, sobre todo, que sean éstos los que clarifiquen sus propias preferencias a fin de facilitar la toma de decisiones vocacionales. La aplicación de la rejilla propuesta puede ser tanto individual como colectiva. En este caso, no obstante, conviene que el grupo no sea excesivamente grande a fin de ofrecer una interpretación lo más individualizada posible.

En esta rejilla, los elementos son salidas profesionales o posibles puestos de trabajo que al sujeto le gustaría desempeñar. Los constructos son características y valores que el sujeto busca realizar a través de una profesión. Las puntuaciones a otorgar a las relaciones elementos/constructos son de intervalo y, en esta ocasión, oscilan en una escala de 1 a 3 puntos.

El proceso de aplicación de la rejilla comenzaría con una introducción general en torno a la actividad que se va a realizar. Los pasos e instrucciones a dar se pueden sintetizar en los siguientes:

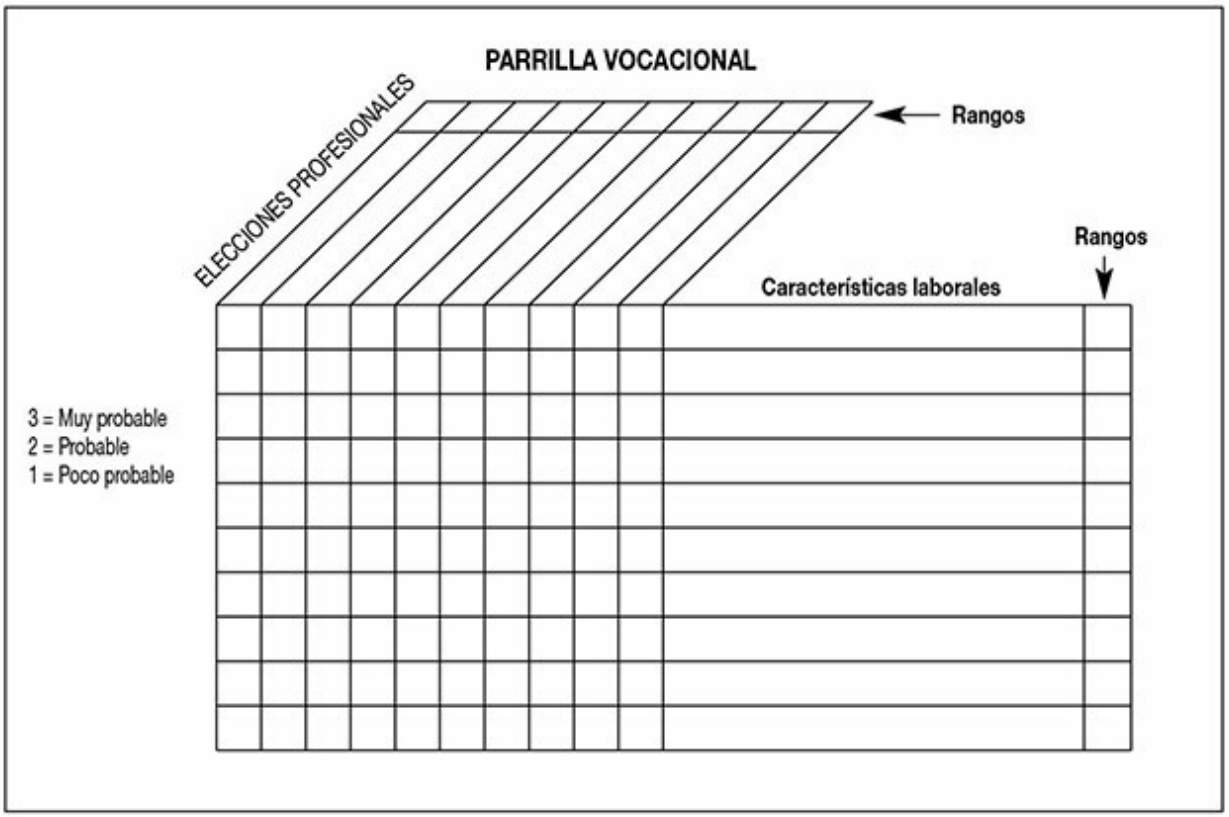
1. Hacer una lista con posibles elecciones profesionales. Es necesario indicar que intente dar hasta 10 actividades profesionales que le gustaría desempeñar, independientemente de su perfil o trayectoria formativa. Cada una de estas profesiones tiene que ser ordenada de 1 a 10 según la preferencia del sujeto, siendo 1 la que más le interesa.
2. Hacer una lista con características, condiciones o valores laborales que la persona considera importantes a la hora de elegir una ocupación (ganar dinero, estabilidad, autonomía, etc.). Nuevamente hay que intentar dar hasta 10 de estas características y, posteriormente ordenarlas de mayor (1) a menor preferencia (10). Como se puede

apreciar, aunque los constructos son aportados por el sujeto, no se elicitan a través del procedimiento de comparación/constraste antes indicado. En aplicaciones individuales del ejemplo, dicho procedimiento se podría igualmente utilizar.

3. Ambos conjuntos de información se disponen en una matriz como la que se representa en el cuadro 10. Ni las profesiones ni las características laborales tienen que estar ordenadas de mayor a menor preferencia, ya que basta con poner en cada una de estas dos variables el rango con que el sujeto las ha puntuado (siendo 1 la más importante).
4. Empezando de izquierda a derecha y de arriba a abajo, el sujeto tiene que puntuar cada ocupación según el grado en que ésta exprese cada valor o característica laboral de las apuntadas, de acuerdo con la siguiente escala: 1= poco probable, 2= probable, 3= muy probable.

CUADRO 10

EJEMPLO VACIO DE REJILLA APLICADA A LA DECISIÓN VOCACIONAL



Tomado de MAU, 1997 y PADILLA, 2001.

En el cuadro 11 se muestra un ejemplo ya cumplimentado. En él se puede observar que el sujeto ha seleccionado 10 actividades profesionales, siendo la de diseñador gráfico, la de adjunto a dirección y la de jefe de administración, las que más le interesan, si atendemos a los rangos que ha asignado en la fila superior a las mismas. Asimismo, ha señalado diez características/valores laborales que le resultan relevantes y significativos. Los que más le atraen son, por este orden, la estabilidad del puesto, el que le proporcione satisfacción personal y la remuneración. Esto se aprecia en la columna denominada «rangos» que aparece junto a las características apuntadas.

CUADRO 11

EJEMPLO DE RESULTADO DE LA REJILLA APLICADA A LA DECISIÓN VOCACIONAL

PARRILLA VOCACIONAL

ELECCIONES PROFESIONALES → / Características laborales	Diseñador gráfico	Jefe de administración	Adjunto de dirección	Consultor RR.HH.	Jefe de personal	Gerente	Asesor empresas	Jefe de organización	Publicidad	Contable	Rangos	Rangos inversos
	1	3	2	5	10	7	4	6	8	9	← Rangos	
Remuneración	2	3	3	2	2	3	2	2	2	1	3	8
Horario flexible	3	1	2	1	1	2	1	1	2	1	4	7
Autonomía	3	3	3	2	3	3	2	2	2	2	5	6
Creatividad	3	1	2	1	1	2	2	3	3	1	7	4
Responsabilidad	1	3	3	2	3	3	2	3	2	1	8	3
Estabilidad	2	3	3	2	2	3	2	2	2	3	1	10
Promoción	1	2	3	1	2	1	1	2	1	2	10	1
Trabajar solo	3	2	2	3	2	2	3	2	3	2	9	2
Tareas variadas	3	3	3	3	2	3	3	2	3	1	6	5
Satisfacción personal	3	3	3	2	1	2	3	3	2	1	2	9
	24	24	27	19	19	24	21	22	22	15	← Suma de columnas	
	3	3	1	8,5	8,5	3	7	5,5	5,5	10	← Rangos	
	139	140	152	105	99	141	118	119	120	84	← Utilidad esperada	
	4	3	1	8	9	2	7	6	5	10	← Rangos	

3 = Muy probable
2 = Probable
1 = Poco probable

Tomado de PADILLA, 2001:165.

Una vez seleccionados y ordenados los elementos (elecciones profesionales) y los constructos (características-valores laborales), el sujeto ha valorado de 1 a 3 el grado en que cada profesión le permite expresar cada valor o característica. Por ejemplo, nuestro sujeto considera como probable que siendo diseñador gráfico pueda tener una buena

remuneración, y como muy probable que, en esta misma profesión, pueda tener un horario flexible y una gran autonomía y desarrollar su creatividad.

A partir de aquí comienza el análisis de la información que puede ser realizado conjuntamente entre el sujeto y el profesional del diagnóstico. Se trata de combinar las informaciones provenientes de las ordenaciones y priorizaciones que ha efectuado el sujeto. Para ello existen dos posibles procedimientos que a continuación explicamos.

El primero de estos procedimientos resulta más sencillo y rápido de calcular. Básicamente consiste en: 1) sumar los valores numéricos de cada columna; 2) asignar un rango a dichos valores (siendo 1 el valor más alto). De esta forma, conocemos cuáles son las ocupaciones que permiten expresar más y mejor cada característica laboral. En el ejemplo del cuadro 11, las profesiones que obtienen un valor más alto según este procedimiento (ver fila denominada «suma de columnas») son «adjunto a dirección», «diseñador gráfico», «jefe de administración» y «gerente». Por esta razón, en la fila inmediatamente inferior se ha asignado a «adjunto a dirección» el rango 1; como las otras tres obtienen la misma puntuación (24), ocuparían los puestos 2, 3 y 4, aunque en este caso de repetición, se pone el valor intermedio entre ellos, o sea 3.

El segundo procedimiento es algo más complejo pero también aporta una información más precisa, ya que en él todas las características laborales no son consideradas al mismo nivel, sino que ponderan según sean más o menos importantes para el sujeto. Para la estimación de las puntuaciones, procedemos del siguiente modo:

1. Para ponderar adecuadamente cada característica laboral, reordenamos los valores inicialmente otorgados a cada una de ellas. Así, la más importante, a la que se le concedió un 1, pasa a valer ahora 10, y la menos importante, a la cual se le otorgó el valor 10, pasa ahora a valer 1. Este cambio de valores aparece recogido en la columna «rangos inversos».
2. Calculamos la utilidad esperada (UE) de cada alternativa profesional elegida. Para ello nos basamos en la siguiente fórmula: UE= (E · V), donde UE es la utilidad esperada, E el valor de 1 a 3 que se ha asignado a cada profesión según cada característica, y V la importancia que cada característica tiene para un sujeto. Tomando como referencia el ejemplo del cuadro 11, en la primera columna el procedimiento se aplicaría así:

   — Tomamos la probabilidad de expresar el valor «remuneración» con la profesión «diseñador gráfico», o lo que es lo mismo, el cruce de la primera fila con la primera columna. El valor numérico en cuestión es 2.

   — Multiplicamos el valor numérico 2 por el rango inverso del valor «remuneración», 8. Obtenemos el valor 16.

   — Repetimos este procedimiento para el resto de valores numéricos de la columna; cada uno de ellos es multiplicado por el rango inverso del valor correspondiente a la fila en la que se encuentran.

— Se suman todos los valores.

— El resumen de los cálculos realizados es: UE= (2 · 8) + (3 · 7) + (3 · 6) + (3 · 4) + (1 · 3) + (2 · 10) + (1 · 1) + (3 · 2) + (3 · 5) + (3 · 9) = 139. Los resultados aparecen en la fila denominada «utilidad esperada».

3. Finalmente a cada valor de UE le asignamos un rango (siendo 1 el rango para el valor numérico más alto). En nuestro ejemplo, el orden de las profesiones es el siguiente: adjunto a dirección (rango 1), gerente (rango 2), jefe de administración (rango 3), diseñador gráfico (rango 4), etc.

De esta forma, obtenemos una ordenación más precisa que por el método anterior. No sólo tenemos en cuenta en qué grado cada actividad profesional permite expresar un determinado valor o característica, sino también si los valores o características que permite expresar son importantes o no para el sujeto.

A pesar de que, a primera vista, puede parecer un procedimiento complejo, este ejemplo de rejilla resulta bastante fácil de aplicar y presenta varias ventajas:

— Por un lado, la información sobre la que se trabaja siempre es proporcionada por el sujeto. Es decir, no se trata de pedirle a éste que valore las profesiones o características que el profesional del diagnóstico le entrega (como sucedería si aplicamos un test de intereses profesionales o un test de valores laborales), sino que, en primer lugar, el sujeto ha de reflexionar sobre las actividades profesionales que más le atraen y cómo considera que cada una de ellas expresa un determinado valor o característica. Por ello se trata de una construcción personal, ya que los valores de 1 a 3 son subjetivos, basados en las percepciones que cada sujeto tiene de una ocupación.

— En segundo lugar, esta rejilla puede ser utilizada no sólo como medio de diagnóstico, sino también como parte de un proceso de intervención. En el ejemplo de ÁLVAREZ ROJO (1997), las profesiones que sirven como elementos de las rejillas fueron seleccionadas por los sujetos en una actividad previa (tarjetas vocacionales) y los valores laborales que actúan como constructos también. La reflexión que, a través de la rejilla, el sujeto tiene que hacer en torno al binomio ocupación-característica, puede también ser aprovechada por el profesional del diagnóstico para plantear nuevas actividades más específicas de exploración. En el caso que hemos puesto como ejemplo (cuadro 11), habría que preguntar al sujeto por qué ha asignado a la ocupación de diseñador gráfico el primer puesto cuando él considera que no le va a permitir expresar la mayoría de los valores y características laborales que persigue desarrollar.

— En tercer lugar, dado que este ejemplo de rejilla se puede realizar en pequeños grupos, es posible planificar otras actividades grupales que sirvan de apoyo al proceso de diagnóstico e intervención. Por ejemplo, los sujetos que hayan elegido las mismas ocupaciones pueden comparar sus resultados en función de los valores

que cada uno busca desarrollar; esta comparación también se puede realizar comparando los resultados de los sujetos que buscan los mismos valores pero han elegido distintas ocupaciones, etc. Todas estas actividades facilitarán una mayor integración psicológica del proceso seguido.

— Por último, ofrece una información simplificada con la que trabajar. Aunque está referida a un número, expresa el resultado de las múltiples decisiones que ha sido necesario tomar para su cumplimentación.

Por otra parte, también en este ejemplo de rejilla se pueden realizar análisis como los mencionados en el apartado anterior. Asimismo, para aquellos profesionales acostumbrados a trabajar con hojas de cálculo (Excell, Quatro-Pro...), el procedimiento de cálculo de las utilidades esperadas puede verse simplificado elaborando el correspondiente programa, en el que se determinan previamente las operaciones a realizar. De esta forma, la obtención de resultados es inmediata una vez introducidos los correspondientes datos. La desventaja de este procedimiento es que el sujeto no participa en el análisis, lo cual es muy importante para que pueda comprender qué significan los resultados. En este sentido, más que el número final, lo que interesa es comprender cómo éste se compone de las valoraciones que él ha hecho sobre la relación ocupación-características laborales que busca desarrollar.

Este formato de rejilla que, como hemos dicho, permite una aplicación colectiva, puede ser utilizado para otros fines, con otros temas y en diversos contextos. Por ejemplo, se nos ocurre que podría ser valioso para evaluar las relaciones interpersonales en ambientes laborales; diferentes estilos docentes en un equipo de profesores; obstáculos y beneficios percibidos en el desempeño de múltiples roles; etc.

Como se ha indicado en numerosas ocasiones, la técnica de rejilla de constructos personales se caracteriza por su versatilidad, por lo que los posibles contextos y situaciones de aplicación dependen de la creatividad del profesional que las usa. Claro que, un buen nivel de dominio de la técnica y una adecuada planificación, son otras condiciones necesarias para conseguir mayor calidad en la aplicación.

Referencias bibliográficas

ALBERS-EBER, C. (2000). *Les enfants placés et la construction d'historicité*. Paris: L'Harmattan.

ANGUERA, M. T. (1995). Recogida de datos cualitativos. En M.T. Anguera et al., *Métodos de investigación en psicología* (pp. 522-547). Madrid: Síntesis.

BERTAUX, D. (1989). Les récits de vie comme forme d'expression, comme approche et comme mouvement. En G. Pineau y G. Jobert (coord.), *Histoires de vie. Utilisation pour la formation.* Tome 1(pp. 17-38). Paris: L'Harmattan.

BERTRAND, J. T., BROWN, J. E. y WARD, V. M. (1992). Techniques for analyzing focus group data. *Evaluation Review, 16* (2), 198-209.

BOTELLA, L. y FEIXAS, G. (1990). El grupo autobiográfico: Un instrumento para la reconstrucción de la experiencia con personas de edad avanzada. *Anuario de Psicología, 44*, 47-60.

BOTELLA, L. y FEIXAS, G. (1993). The autobiographical group: A tool for the reconstruction of past life experience with the aged. *International Journal of Aging and Human Development, 4*, 303-319.

BOTELLA, L. y FEIXAS, G. (1998). *Teoría de los constructos personales: Aplicaciones a la práctica psicológica*. Barcelona: Laertes.

CANALES, M. y PEINADO, A. (1994). Grupos de discusión. En J.M. Delgado y J. Gutiérrez (coord.), *Métodos y técnicas cualitativas de investigación en ciencias sociales* (pp. 287-316). Madrid: Síntesis.

CAPRA, F. (1998). *La trama de la vida. Una nueva perspectiva de los sistemas vivos*. Barcelona: Anagrama.

CENTRO DE ADULTOS «JUAN XXIII» (1999). *Trozos de mi vida*. Sevilla: Fundación El Monte.

COLE, D. J., RYAN, CH. W. y KICK, F. (1995). *Portfolios across the curriculum and beyond*. Thousand Oaks, CA: Corwin Press.

DELBECQ, A., VAN DE VEN, A. y GUSTAFSON, D. (1984) *Técnicas grupales para la planeación*. México, DF: Trillas.

DELORY-MOMBERGER, C. (2000). *Les histoires de vie. De l'invention de soi au projet de formation.* Paris: Anthropos.

DE MIGUEL, J. M. (1996). *Auto/biografías*. Madrid: Centro de Investigaciones Sociológicas.

FEIXAS, G. y CORNEJO, J. J. (1996). *Manual de la técnica de rejilla mediante el programa RECORD v. 2.0.* Barcelona: Paidós.

FEIXAS, G. y VILLEGAS, M. (1990). Evaluación de textos autobiográficos. Aplicación de un diseño metodológico basado en la teoría de los constructos personales de Kelly: El caso Pablo Neruda. *Evaluación Psicológica, 6* (3), 289-326.

FERNÁNDEZ BALLESTEROS, R. (1992). Técnicas subjetivas. En R. Fernández Ballesteros (coord.), *Introducción a la evaluación psicológica* (pp. 279-313). Madrid: Pirámide.

FORTIN, R. (2000). *Comprende la complexité. Introduction à la Métode d'Edgar Morin*. Laval: L'Harmattan.

FRANSELLA, F. y BANNISTER, D. (1977). *A manual for repertory grid technique*. London: Academic Press.

GONZÁLEZ MONTEAGUDO, J. (1996). Historias de vida. Aspectos históricos, teóricos y epistemológicos. *Cuestiones Pedagógicas, 12*, 223-242.

HERMAN, J. L., GEARHART, M. y ASCHBACHER, P. R. (1996). Porfolios for classroom assessment: Design and implementation issues. En R. Calfee y P. Perfumo (eds.), *Writing portfolios in the classroom. Policy an practice, promise and peril* (pp. 27-59). Matwath, NJ: Lawrence Erlbaum Associates.

HORMIGO GONZÁLEZ, J. (2000). *Tiempos difíciles. Memorias de un trabajador*. Sevilla: Imprenta Galán.

KELLY, G. A. (1955, 1991). *The psychology of personal constructs* (vols. 1 y 2). London: Routledge.

KRUEGER, R. A. (1991). *El grupo de discusión. Guía práctica para la investigación aplicada*. Madrid: Pirámide.

LERAY, C. y BOUCHARD, C. (2000). *Histoire de vie et dynamica langagière*. Rennes: Presses Universotaires de Rennes.

LIETARD, B. (1998). Accompagner, oui. Mais comment? En G. Pineau. (ed.), *Accompagnements et histoire de vie* (pp. 113-118). Paris: L'Harmattan.

MAU, W. (1997). Assessing the cognitive complexity of vocational value constructs using the career grid: a comparison of the expressed and supplied methods. *Measurement and Evaluation in Counseling and Development, 29*, 202-214.

MORENO, E.; ARANDA, J. M. y MONTERO, M. (1993). *Determinantes cognitivos implicados en el uso de los métodos anticonceptivos en las mujeres andaluzas.* Inédito. Investigación subvencionada por el Instituto Andaluz de la Mujer.

MUCCHIELLI, R. (1978). *La entrevista en grupo*. Bilbao: Mensajero.

ORTI, A. (1986). La apertura y el enfoque cualitativo o estructural: la entrevista abierta semidirectiva y la discusión de grupo. En M. García Ferrando, J. Ibáñez y F. Alvira (comp.), *El análisis de la realidad social. Métodos y técnicas de investigación* (pp. 153-185). Madrid: Alianza Editorial.

PADILLA, M. T. (2001). *Estrategias para el diagnóstico y la orientación profesional de personas adultas.* Barcelona: Laertes.

PÉREZ-CAMPANERO, M. P. (1991). *Cómo detectar las necesidades de intervención socioeducativa*. Madrid: Narcea.

PÉREZ-PAREJA, F .J. (1997). Autoinformes. En G. Buela-Casal y J.C. Sierra (dirs.), *Manual de evaluación psicológica* (pp. 296-314). Madrid: Siglo Veintiuno Editores.

PINEAU, G. y MARIE-MICHÈLE (1983). *Produire sa vie: Autoformation et autobiographie.* Québec:Éditions S. Martin.

PINEAU, G. y JOBERT, G. (coord.) (1989). *Histoires de vie. Utilisation pour la formation.* Tome 1. Paris: L'Harmattan.

PINEAU, G. y JOBERT, G. (coord.) (1989). *Histoires de vie. Approches multidisciplinaires*. Tome 2. Paris: L'Harmattan.

PINEAU, G. (1998). *Accompagnements et histoire de vie.* Paris: L'Harmattan.

PORTER, C. y CLELAND, J. (1995). *The portfolio as learning strategy*. Portsmouth, NH: Boyton/Cook.

RIVAS, F. y MARCO, R. (1985). *Evaluación conductual subjetiva: la técnica de rejilla*. Valencia: Centro Editorial de Publicaciones y Servicios Universitarios.

RODRÍGUEZ ESPINAR, S. (1997). El portafolios: ¿modelo de evaluación o simple historial del alumno? En H. Salmerón Pérez (coord.). *Diagnosticar en educación. Actas de las III Jornadas Andaluzas de Orientación* (pp. 183-199). Granada: FETE-UGT.

SALOMÉ, J. y GALLAND, S. (1990). *Si je m'écoutais, je m'entendrais*. Québec: Les Éditions de l'Homme.

SANTAMARINA, C. y MARINAS, J. M. (1994). Historias de vida e historia oral. En J.M. Delgado y J. Gutiérrez (coord.), *Métodos y técnicas cualitativas de investigación en ciencias sociales* (pp. 257-285). Madrid: Síntesis.

WARD, V. M., BERTRAND, J. T. y BROWN, L. F. (1991). The comparability of focus group and survey results. *Evaluation Review, 15* (2), 266-283.

WITKIN, B. R. y ALTSCHULD, J. W. (1995). *Planning and conducting needs assessments. A practical guide*. Thousands Oaks, CA: Sage.

Anexos

Anexo 1
CÓDIGO DEONTOLÓGICO PROFESIONAL

Código Deontológico del Profesional de la Pedagogía aprobado por la asamblea de Pedagogía en el III Encuentro Estatal de Pedagogos (Valdepeñas, Ciudad Real, 1986)

Introducción

La profesión del Pedagogo/a, al igual que otras profesiones que implican el trato directo con la persona y cuyo campo de trabajo es la sociedad, debe tener en cuenta en su desarrollo ciertas normas a que atenerse. Estas normas deontológicas son necesarias e imprescindibles para el buen funcionamiento y normalización de la profesión, independientemente de la labor y campo en que se aplique su conocimiento.

Normas deontológicas

1. El/la Pedagogo/a no debe colaborar en trabajos y estudios que vayan en contra de la Declaración Universal de los Derechos del Hombre, ni en procesos que atenten contra la dignidad del ser humano y su capacidad de elección libre y racional.

2. El profesional deberá tener como actitud principal una buena atención a la persona y un adecuado servicio a la sociedad, teniendo en cuenta que esa relación no debe conducir a una dependencia que provoque un estado de indefensión del niño, del adulto o de la comunidad.

3. El/la Pedagogo/a velará para que las técnicas de intervención potencien el crecimiento y la comprensión de la realidad, tendiendo al desarrollo integral de la persona.

4. El/la Pedagogo/a ha de respetar la autonomía, libertad de decisión y dignidad de otro profesional y evitar la mutua desautorización entre ellos porque defiendan teorías pedagógicas o jerarquías de valores diferentes.

En beneficio de la profesión, se debe evitar toda forma de competencia desleal.

5. El/la Pedagogo/a deberá mantener una formación continua para un buen desempeño de su tarea profesional.

La profesión del Pedagogo/a puede ejercerse de varias formas: a) Forma liberal. b) Mediante relación laboral con la empresa: pública y privada. c) Mediante relación funcionarial.

6. El/la Pedagogo/a deberá difundir por los medios posibles, y en su caso hacer partícipe al Colegio Profesional, aquellos programas, técnicas o métodos innovadores que puedan contribuir a la mejora de la profesión.

7. El/la Pedagogo/a deberá actuar con la máxima objetividad, respetando y guardando el secreto profesional.

8. El/la Pedagogo/a trabajará de manera que puedan analizarse las consecuencias de su actuación. Deberá tener en cuenta que su inhibición en determinados momentos y aspectos del ejercicio de su profesión puede ser perjudicial.

9. Actuará el/la Pedagogo/a solidariamente en defensa de su profesión, siempre que no vaya en contra de este código deontológico.

10. Estará abierto el/la Pedagogo/a a un proceso de colaboración e intercambio con otros profesionales.

Anexo 2
PROPUESTA DE NORMAS ÉTICAS PARA LA ELABORACIÓN DE UN CÓDIGO DEONTOLÓGICO DEL PROFESIONAL DE LA PEDAGOGÍA (Versión 2001)

El primer borrador de este documento fue elaborado a partir de la revisión de otros códigos deontológicos de profesiones afines a la Pedagogía (AIOSP, 1995; COP, 1987; NCDA, 1987), por lo que algunos de sus artículos son muy similares a los contenidos en los mismos. En la elaboración de este documento han participado otras personas además de la autora del libro. Entre ellas, cabe destacar las valiosas aportaciones de diferentes profesionales en ejercicio de la Pedagogía, así como de las siguientes promociones de estudiantes de pedagogía: estudiantes de *Diagnóstico Pedagógico* curso 1998/99 (todos los grupos), estudiantes de *Diagnóstico Pedagógico* curso 1999/00 (grupos de tarde), estudiantes de *Técnicas e Instrumentos de Diagnóstico* curso 2000/01 (grupos P2 y P3). El resultado final que aquí se presenta intenta recoger todas las aportaciones realizadas por los colectivos mencionados, si bien, esta propuesta sigue siendo sometida al debate y la reflexión.

PROPUESTA DE NORMAS ÉTICAS PARA LA REGULACIÓN DE LA CONDUCTA PROFESIONAL DEL PEDAGOGO (VERSIÓN 2001)

Art. 1 El ejercicio de la Pedagogía se orienta a una finalidad humana, social, educativa y cultural que puede expresarse en objetivos tales como: el bienestar, la calidad de vida y de enseñanza, la plenitud del desarrollo de las personas y de los grupos, procurando siempre la integración de la persona en el proceso educativo y en la sociedad en general.

Art. 2 Puesto que el/la Pedagogo/a no es el/la único/a profesional que persigue los fines enunciados en el artículo anterior, es conveniente, enriquecedora y necesaria, en muchos casos, la colaboración interdisciplinar con otros/as profesionales, sin perjuicio de las competencias y saberes de cada uno/a de ellos/as.

Art. 3 La profesión de Pedagogo/a se rige por principios comunes a toda deontología profesional: respeto a la persona, protección de los derechos humanos, sentido de responsabilidad, honestidad y claridad para con las personas destinatarias de la intervención, prudencia en la aplicación de instrumentos y técnicas, competencia profesional, solidez de la fundamentación objetiva y científica de sus intervenciones profesionales.

Art. 4 Todos los profesionales de la Pedagogía, dentro de sus posibilidades de actuación, tienen la obligación y el deber de dar respuestas a cualquier tipo de necesidad educativa manifestada en su intervención, respetando en todo momento las diferencias individuales y grupales en los procesos de intervención pedagógica.

Art. 5 El/la profesional de la Pedagogía es un agente de cambio educativo y social y, como tal, ha de promover aquellos procesos de innovación socio-educativa que se estimen necesarios.

Art. 6 En un proceso de evaluación y/o de intervención educativa, el/la profesional de la Pedagogía procurará adoptar una actitud objetiva, no permitiendo que sus opiniones o sentimientos personales puedan condicionar el proceso y los resultados.

Art. 7 El/la Pedagogo/a actuará en todo momento con honestidad y no aprovechará para lucro o beneficio propio o de terceros, la situación de poder o superioridad que el ejercicio de la profesión pueda conferirle sobre las personas destinatarias de su intervención.

Art. 8 En sus experiencias laborales, y especialmente en sus informes escritos, el/la Pedagogo/a será sumamente cauto, prudente y crítico frente a nociones que fácilmente degeneran en etiquetas devaluadoras y discriminatorias, del género de normal/anormal, adaptado/inadaptado, inteligente/deficiente, u otras similares.

Art. 9 La autoridad profesional del Pedagogo/a se fundamenta en su capacitación y cualificación para las tareas que desempeña. El/la Pedagogo/a ha de estar profesionalmente preparado y especializado en la utilización de métodos, instrumentos, técnicas y procedimientos que adopte en su trabajo. Forma parte de su trabajo el esfuerzo continuado de actualización de sus competencias profesionales. Debe reconocer los límites de su competencia y las limitaciones de sus técnicas, intentando ampliarlos y mejorarlos en la medida de sus posibilidades.

Art. 10 Todo material estrictamente psicopedagógico, tanto de evaluación cuanto de intervención o tratamiento, queda reservado al uso de los/las Pedagogos/as, quienes, por otra parte, se abstendrán de facilitarlos a otras personas no competentes. No obstante, y puesto que toda intervención educativa precisa de la colaboración de diversos agentes, el/la Pedagogo/a garantizará la custodia y uso legítimo de tales documentos, siendo conscientes del grado de responsabilidad que conlleva su uso y manipulación.

Art. 11 El/la Pedagogo/a debe tener especial cuidado en no crear falsas expectativas que después sea incapaz de satisfacer profesionalmente, manteniendo siempre una actitud positiva ante la intervención que puede realizar. Su labor estará basada en el análisis de necesidades y la planificación de los medios técnicos y recursos adecuados para la satisfacción de las mismas.

Art. 12 Toda información que el/la Pedagogo/a recoge en el ejercicio de su profesión, sea en manifestaciones verbales expresas de las personas con las que interviene, sea en

datos psicotécnicos o en otras observaciones profesionales practicadas, está sujeta a un deber y a un derecho de secreto profesional del que sólo podría ser eximido por el consentimiento expreso de la persona afectada por su intervención, o de los representantes legales de ésta cuando se trate de un/a menor de edad. No obstante, esta información podrá ser compartida con otros agentes educativos cuya opinión y colaboración se requiera con fines de mejora. En estos casos, el/la Pedagogo/a velará porque sus eventuales colaboradores se atengan a este secreto profesional.

12.1. Cuando la evaluación o intervención psicopedagógica se produce a petición del propio sujeto de quien el/la Pedagogo/a obtiene información, ésta sólo puede comunicarse a terceras personas, con expresa autorización previa del interesado o de su representante legal y dentro de los límites de esta autorización.

12.2. Cuando dicha evaluación o intervención haya sido solicitada por otras personas ajenas al proceso de intervención educativa —jueces, empleadores, o cualquier otro solicitante diferente del sujeto evaluado—, éste último o sus padres o tutores tendrán derecho a ser informados del hecho de la evaluación o intervención y del destinatario del informe pedagógico consiguiente. El sujeto de un informe pedagógico tiene derecho a conocer el contenido del mismo, siempre que de ello no se derive un grave perjuicio para él/ella o para el/la Pedagogo/a, y aunque la solicitud de su realización haya sido hecha por otras personas.

12.3. Los informes pedagógicos realizados a petición de instituciones u organizaciones públicas o privadas, aparte de lo indicado en los artículos anteriores, estarán sometidos al mismo deber y derecho general de confidencialidad antes establecido, quedando tanto el/la Pedagogo/a como la correspondiente instancia solicitante obligados a no darles difusión fuera del estricto marco para el que fueron recabados.

Las enumeraciones o listas de sujetos evaluados en los que deban constar los diagnósticos o datos de la evaluación y que se les requieran al Pedagogo/a por otras instancias, a efectos de planificación, obtención de recursos u otros motivos, deberán realizarse omitiendo el nombre y datos de identificación del sujeto, siempre que no sean estrictamente necesarios.

12.4. La exposición oral, impresa, audiovisual u otra, de casos de intervención educativa con fines didácticos o de comunicación o divulgación científica, debe hacerse de modo que no sea posible la identificación de la persona, grupo o institución de que se trata.

En el caso de que el medio usado para tales exposiciones conlleve la posibilidad de identificación del sujeto, será necesario su consentimiento previo explícito.

Art. 13 La evaluación y la intervención pedagógica deberán tener en cuenta los contextos en los que interactúa la/s persona/s destinataria/s de la actuación. Como

consecuencia de ello y en la medida de lo posible, estos contextos serán también objeto de evaluación e intervención del profesional de la Pedagogía.

Art. 14 El/la Pedagogo/a debe ofrecer orientación e información específica a la persona/institución destinataria de un proceso de evaluación antes y después de la administración de los oportunos instrumentos y técnicas de diagnóstico. Explicará el propósito de la evaluación y el uso que se va a hacer de ella y sus resultados.

Art. 15 En la selección de instrumentos o técnicas para una situación de evaluación dada, el/la Pedagogo/a evaluará cuidadosamente el modelo teórico desde el cual está construido cada instrumento, así como su validez, fiabilidad y adecuación para el caso a evaluar. El profesional de la Pedagogía será responsable del uso de información no validada empíricamente.

Art. 16 En un proceso diagnóstico, el/la Pedagogo/a vigilará que se cumplan las condiciones para las cuales están diseñados los instrumentos de medida, procurando que durante su aplicación no tengan lugar comportamientos inusuales o irregularidades, y cuestionando la validez de los resultados cuando así sucediera.

Art. 17 El/la Pedagogo/a procederá con precauciones especiales cuando realice evaluaciones de grupos minoritarios u otras personas que no están representadas en el grupo normativo con el que se validó y estandarizó el instrumento usado.

Art. 18 La comunidad pedagógica debe velar por la continua creación, mejora y adaptación de las técnicas e instrumentos pedagógicos, a fin de enriquecer y diversificar los procedimientos de evaluación e intervención de los/as profesionales de la Pedagogía.

Art. 19 Los informes pedagógicos habrán de ser claros, precisos, rigurosos e inteligibles para su destinatario, ofreciendo siempre respuestas de intervención educativa individualizadas para cada situación concreta. Deberán expresar su alcance y limitaciones, el grado de certidumbre que acerca de sus varios contenidos posea el informante, su carácter actual o temporal, las técnicas utilizadas para su elaboración, haciendo constar en todo caso los datos del profesional que lo emite.

Anexo 3
EJEMPLOS DE SISTEMAS CATEGORIALES DE OBSERVACIÓN

1) EJEMPLO DE SISTEMA DE CATEGORÍAS

Este ejemplo ha sido extraído de FLANDERS (1977)[37]. Aquí se recogen sólo las categorías del sistema, pero éstas derivan de un marco teórico expuesto en la obra citada. Los intervalos de observación del sistema son de tres segundos de duración. En la hoja de registro, además de unos datos situacionales (grupo, profesor, fecha...), en los recuadros correspondientes a cada intervalo se consigna el número de la/s categoría/s que representan la/s conducta/s observadas en el intervalo.

PARTICIPACIÓN ORAL DEL PROFESOR	**INFLUENCIA DIRECTA**	1. **Acepta los sentimientos.** Acepta y aclara los sentimientos del alumno de forma no amenazadora. Los sentimientos pueden ser positivos y negativos. Se incluyen los sentimientos predictivos y los de recuerdo.
		2. **Alaba o anima la acción o comportamiento del alumno.** Chistes para relajar la tensión, no a costa de otro individuo; se incluyen los asentimientos de cabeza y frases como «bien», «sí».
		3. **Acepta o usa ideas del estudiante.** Clarificación, construcción o desarrollo de ideas o sugerencias de un estudiante. Si el profesor pone en juego más ideas propias pasa a la categoría 5.
	INFLUENCIA INDIRECTA	4. **Hace preguntas** sobre el contenido o el método con la intención de que el alumno responda.
		5. **Instruye** dando hechos u opiniones sobre el contenido o los métodos expresando sus propias ideas; haciendo preguntas retóricas.
		6. **Da instrucciones,** direcciones, órdenes o mandatos que debe cumplir el estudiante.
		7. **Critica o justifica la autoridad.** Afirmaciones para cambiar el comportamiento del estudiante de una pauta no aceptable a otra aceptable; echar una bronca a alguien; afirmando porqué el profesor está haciendo lo que hace, referencia excepcional a sí mismo.

PARTICIPACIÓN ORAL DEL ESTUDIANTE	8. **Respuesta oral del estudiante.** Los estudiantes hablan en respuesta al profesor. El profesor inicia el contacto o solicita una afirmación del estudiante. 9. **El estudiante inicia la charla.** Charla de los estudiantes, iniciada por ellos. Si se nombra a un estudiante solamente para indicar quién puede hablar a continuación, el observador debe decidir si el estudiante quería hablar. En caso afirmativo, usa esta categoría. 10. **Silencio o confusión.** Pausas, períodos cortos de silencio y períodos de confusión en los que el observador no puede entender la comunicación.

2) EJEMPLO DE LISTA DE CONTROL

El instrumento que presentamos es una lista de control que no está dirigida a la observación del comportamiento de sujetos, sino que se enfoca a la observación del tratamiento de las diferencias de género en los planes de orientación y acción tutorial de centros de enseñanza primaria y secundaria

CENTRO: OBSERVADOR: FECHA:		
	SÍ	**NO**
— Se abordan explícitamente las diferencias de género		
— Se plantea entre los objetivos la potenciación de una orientación para la igualdad de géneros		
— Se plantean actividades orientadoras para favorecer una elección profesional no sexista		
— Se plantean actividades orientadoras para trabajar diferencialmente el autoconcepto y la autoestima		
— Se plantean actuaciones coordinadas para favorecer una educación no sexista		

3) EJEMPLO DE ESCALA DE ESTIMACIÓN

Esta escala de estimación permite observar el comportamiento de sujetos ante el trabajo en grupo. Se puede aplicar en situaciones que requieran la colaboración de un grupo para resolver una tarea.

SUJETO:
OBSERVADOR:
FECHA:

	NUNCA	POCAS VECES	ALGUNAS VECES	MUCHAS VECES	SIEMPRE
Respeta el turno de palabra					
Permanece en el grupo mientras dura la tarea					
Hace comentarios despectivos ante ideas y opiniones distintas a las suyas					
Inicia conversaciones o realiza comentarios sobre temas colaterales a la tarea					
Su tono de voz cuando se expresa es conversacional (no grita o habla en voz alta para remarcar su opinión)					
Respeta las normas de funcionamiento que ha consensuado el grupo					
Adopta una postura corporal adecuada a la situación (no da la espalda a sus compañeros, ni denota pasividad en su postura)					
Toma notas o registra las aportaciones que se van realizando					
Hace comentarios que organizan el trabajo del grupo					
Hace intervenciones que facilitan la realización de la tarea					

Anexo 4
EJEMPLOS DE APLICACIONES DE LAS TÉCNICAS DE ENCUESTA

1) EJEMPLO DE GUIÓN PARA UNA ENTREVISTA ABIERTA SOBRE EL CLIMA DE RELACIONES EN UNA EMPRESA

NOTA: Antes de transcribir el ejemplo, queremos recordar que el guión sólo sirve a la finalidad de orientar la entrevista abierta. Las temáticas en ella abordadas sólo son propuestas, ya que durante el transcurso de la entrevista pueden explorarse otros temas que aparezcan como relevantes. Lo mismo sucede con las preguntas; en el guión que presentamos, las preguntas propuestas sólo están pensadas para generar abundante información y para que sirvan de ejemplo. Ni tienen por qué formularse exactamente igual, ni en el orden y secuencia que figura en el guión.

Objetivos

— Conocer cómo es el clima de relaciones personales en la empresa X y qué opinan y sienten los trabajadores encuestados respecto al mismo.

— Describir y profundizar en los distintos tipos de relaciones que se establecen entre los trabajadores y entre éstos y sus superiores.

— Comprender las formas y mecanismos que las personas que trabajan en la empresa ponen en práctica para relacionarse con sus compañeros y superiores

Destinatarios y selección

Una selección de los trabajadores de la empresa X. En la selección final de entrevistados habrá que procurar que exista representación de trabajadores de todos los departamentos de la empresa. Asimismo se preferirán a aquellos sujetos que denoten una mayor disposición a hablar y a contar su experiencia al entrevistador. Será necesario entrevistar a uno o varios de los trabajadores que son vistos por los demás como más «problemáticos» en las relaciones, así como a los posibles líderes o trabajadores admirados por su forma de relacionarse con los demás.

Temáticas a explorar

Ambiente general de la empresa en la rutina habitual de trabajo; percepción de las relaciones formales e informales entre los departamentos de la empresa; percepción de las relaciones formales e informales entre compañeros del mismo departamento; percepción de las relaciones formales e informales con los superiores.

Guión

1. Ambiente general de la empresa y rutinas de trabajo:

Se trata de conocer con detalle cuál es la rutina habitual de cada entrevistado; qué hace día a día, con quién se relaciona para cada tarea y cómo; cómo percibe la empresa en la que trabaja, qué piensa del ambiente que hay en ella... Esta información permitirá entablar una relación relajada para la entrevista, así como conocer el marco general desde el que cada entrevistado se relaciona con sus compañeros.

Posibles preguntas:

— ¿Puedes describir lo que haces en un día normal de trabajo?

— ¿Cómo te sientes con el trabajo que realizas?

— Si llegara un trabajador nuevo a esta empresa, ¿cómo crees que se sentiría los primeros días de trabajo?

— ...

2. Relaciones entre los departamentos de la empresa:

Se trata de conocer cómo se perciben las relaciones entre los distintos departamentos o secciones de la empresa. ¿Cuál es el clima general de un determinado departamento, tal como lo perciben los trabajadores de otro?, ¿qué diferencias existen en las relaciones de un departamento a otro?, ¿qué mecanismos ponen en práctica los trabajadores para relacionarse con el personal de otros departamentos?...

Posibles preguntas:

— Respecto a la empresa en su conjunto, con sus distintos departamentos, ¿crees que la gente se siente a gusto trabajando aquí?

— Si mañana te cambiaran de departamento, ¿a cuál te gustaría que te destinaran?, ¿por qué?, ¿en cuál no te gustaría trabajar?, ¿por qué?, etc.

— Cuando el trabajo que estás haciendo requiere que vayas de visita al departamento Y, ¿qué sientes?, ¿qué reacciones te provoca esta situación?

— ...

3. Relaciones entre compañeros del mismo departamento:

Se persigue ahora conocer cómo se relaciona el entrevistado con los compañeros con que trabaja en su contexto mas inmediato, codo con codo. ¿Cuál es el clima de trabajo y relaciones que se percibe en el departamento?, ¿cómo son las relaciones informales?, ¿qué nivel de satisfacción existe respecto al clima relacional?...

Posibles preguntas:

— Supón que acabas de casarte y es el primer día de trabajo después de tu viaje de novios,¿qué pasa cuando llegas a la oficina?

— ¿Con qué compañeros te encuentras más a gusto?, ¿a qué crees que se debe esto?...

— Recuerda alguna situación concreta en que hayas tenido un problema personal grave (la enfermedad de un familiar, un problema financiero...), ¿cómo se comportaban tus compañeros?, ¿qué hacían?

— ....

4. Relaciones con los superiores:

A través de este tema se quiere explorar cómo son las relaciones formales o informales que los entrevistados mantienen con sus superiores y cómo son percibidas y valoradas por aquéllos. ¿Están estas relaciones cargadas de tensión?, ¿son excesivamente formalistas?, ¿existen puntos conflictivos?...

Posibles preguntas:

— ¿Podrías describir las relaciones que mantienes con tu jefe o jefa?

— Cuando tienes un problema personal con un compañero, ¿hablas con tu jefe/a?, ¿cómo se lo planteas?...

— ¿Cómo serían, en tu opinión, las relaciones ideales con un superior?

— ¿Podrías poner un ejemplo de alguna situación en que la relación con tu jefe/a te pareciera conflictiva o tensa?

— ...

2) EJEMPLO DE PROCESO DE ELABORACIÓN DE UNA ENTREVISTA/CUESTIONARIO

NOTA: El ejemplo que ahora presentamos no está completo. Para ganar en claridad, se ha optado por hacer un desarrollo más esquemático. En él se definen los objetivos generales del instrumento y el marco general desde el que se elabora (aunque para su construcción no se ha realizado una revisión teórica tan exhaustiva como sería deseable). Se explicitan también las dimensiones o constructos que se recogerán en el instrumento, **desarrollando sólo una de ellas y sus diferentes indicadores** («preferencias literarias»).

Objetivos

1. Conocer las actitudes de los adolescentes ante la lectura, determinando qué opinión tienen de esta actividad.

2. Conocer las preferencias literarias de chicos y chicas, determinando tanto el formato (prensa, revistas, libros...), como el contenido (ciencia ficción, historia...) de lo que suelen leer.
3. Indagar cuáles son las pautas temporales en las que los adolescentes llevan a cabo la lectura, lo que implica conocer tanto el tiempo que dedican a esta actividad los chicos y chicas, como el momento en el que la realizan.
4. Describir las formas y procedimientos a través de los cuales los adolescentes acceden a los libros y otro tipo de lecturas, conociendo si compran y coleccionan los materiales que leen, si van o no las bibliotecas, etc.

Marco conceptual

A través de la lectura, los niños aprenden a interpretar el mundo en el que viven, por lo que poseer unos buenos hábitos de lectura contribuirá positivamente al desarrollo personal y social de los estudiantes. En el marco académico, los hábitos lectores también ejercen una poderosa influencia. La investigación ha demostrado que existe una importante relación entre los hábitos lectores y el rendimiento en la escuela. Así, aquellos estudiantes que tienen la lectura entre sus actividades habituales suelen obtener buenas calificaciones. La lectura contribuye a aumentar el vocabulario, fomenta el razonamiento abstracto, potencia el pensamiento creativo, estimula la conciencia crítica, etc. (MINISTERIO DE EDUCACIÓN, CULTURA y DEPORTES, 2001[38]). Además de todo esto, la lectura es una fuente de placer, pudiendo llegar a considerarse un vicio en tanto que acción que satisface una necesidad (ROSINÉS, 1996[39]). Leer no debe ser algo obligatorio, sino que es una actividad que realizamos para satisfacer una necesidad intelectual y sentimental.

Por todas esta razones, es evidente que fomentar los hábitos lectores debe ser una prioridad del sistema educativo. Para ello, en primer lugar, debemos conocer cómo son los hábitos lectores y cuáles son las conductas habituales de niños y adolescentes respecto de la lectura. Se han realizado ya algunos estudios al respecto en el contexto español, llegándose a la conclusión de que los españoles leen poco, además de que la afición a la lectura disminuye con la edad. Así, si bien en los primeros años escolares esta afición es mayor, conforme el niño se convierte en adolescente, va disminuyendo, por lo que puede considerarse la adolescencia como una etapa crítica.

Por otra parte, leer es un proceso complejo que consiste en saber descifrar códigos lingüísticos, comprender lo que se lee, ser capaz de juzgar su contenido y de gustar por la lectura (FOUCAMBERT, 1989[40]). ¿Qué significa, por tanto, «tener hábitos lectores»? En un intento de definición del término, podemos decir que todo hábito es un modo especial de proceder o conducirse, adquirido por repetición de actos iguales o semejantes originando tendencias instintivas (REAL ACADEMIA DE LA LENGUA, 1992[41]). Así, la repetición parece ser un elemento clave en la definición y caracterización del hábito. En lo relativo a la lectura, los hábitos lectores, por tanto, harán referencia al ejercicio o

práctica repetida de la lectura en situaciones diferentes. Esto implica que leer no sólo debe ser una actividad vinculada estrictamente al ámbito escolar, sino que debe convertirse en una actividad habitual para que pueda hablarse de verdadero hábito lector (MORENO, 2000[42]).

Por tanto, la lectura, como hábito, constituye un proceso humano que se caracteriza por un aprendizaje concreto que va desde la adquisición del automatismo hasta el gusto por la lectura, destacando el carácter de aprendizaje personal, sistemático y constante y, aunque se considera una técnica instrumental, su aprendizaje y perfeccionamiento abarca la vida de la persona (MORENO, 2000).

En la formación de estos hábitos lectores, determinados contextos o situaciones van a ejercer una poderosa influencia. El placer por leer no se despierta de forma automática, sino que supone un esforzado aprendizaje en el que resulta necesaria la estimulación ambiental. La escuela es una institución con una poderosa influencia en el aprendizaje del proceso lector, así como uno de los contextos más importantes en el desarrollo de los hábitos lectores, aunque no el único.

Los hábitos lectores, en tanto que actitudes y valores que las personas portan y moldean durante su vida, tienen también su fundamento en la calidad o tipo de ambiente familiar en el que ésta se desarrolla en los primeros años de su vida. RODRÍGUEZ, MORENO y MUÑOZ (1988[43]) analizaron la relación existente entre diversos aspectos del entorno familiar y medidas globales de la habilidad lectora, demostrando la incidencia de factores como el tipo de interacciones que mantienen los padres con los hijos, la cantidad de tiempo que pasan juntos, los propios hábitos lectores de los padres, etc. Algunos estudios (ARAGÓN, GÓMEZ y JUAN, 1997[44]; MORENO, PADILLA y VÉLEZ, 1998[45]) han puesto en evidencia que la influencia que ejercen los padres en la motivación hacia la lectura es aún mayor que la ejercida por los profesores en el contexto escolar.

Teniendo en cuenta todas estas aportaciones, a la hora de construir un instrumento que pretende evaluar los hábitos lectores de los adolescentes, consideramos que éste debe recoger información sobre las siguientes **dimensiones:**

— La opinión y actitudes que los encuestados tienen ante la actividad lectora y la importancia que conceden a esta actividad, frente a otras formas de pasar el tiempo libre.

— Las preferencias literarias que tienen: el tipo de lecturas que realizan (diarios, revistas, libros, multimedias...) y los contenidos o temas que prefieren leer.

— El tiempo y momento/s que dedican a la lectura.

— La forma en que acceden a los libros. Es decir, en qué medida compran y coleccionan libros y/o asisten a las bibliotecas con asiduidad.

Todas estas dimensiones serán incluidas en el instrumento que pretendemos elaborar, aunque existen otros factores, como la ya apuntada influencia del contexto escolar y

familiar, que, a pesar de estar muy relacionados con los hábitos lectores, no serán tenidos en cuenta. En este sentido, a través de nuestro instrumento pretendemos indagar en las conductas y actitudes de los adolescentes ante la lectura, por lo que entendemos que los contextos escolar y familiar, aunque pueden influir en los hábitos lectores, no constituyen hábitos lectores en sí mismos.

Definición de la dimensión «Preferencias literarias» y presentación de sus indicadores:

Las preferencias literarias se refieren tanto a los temas que los adolescentes eligen para las lecturas que realizan (género literario), como al formato concreto en que se presentan esas lecturas. Así, los hábitos lectores de un adolescente pueden concentrarse en un conjunto de temáticas que captan su interés (la aventura, el deporte, el humor, el romanticismo...) y pueden realizarse, bien a través de la lectura de libros, o bien a través de la lectura de otro tipo de materiales como son los periódicos, las revistas y los cómics.

Algunas investigaciones (ARAGÓN, GÓMEZ y JUAN, 1997; MORENO, 2000; MORENO, PADILLA y VÉLEZ, 1998; etc.) han puesto de manifiesto que, entre las lecturas preferidas por los adolescentes están los cómics, tebeos y revistas, mostrando este colectivo también especial preferencia por los libros de aventura y fantasía. Otros tipos de contenidos (como la poesía, el teatro, la historia, etc.) suelen ser tratados con mayor indiferencia.

Teniendo esto en cuenta, los **indicadores** que consideramos más relevantes para esta dimensión son:

— Tipos de lectura que realizan los adolescentes. Se refiere al material impreso que este colectivo suele leer preferentemente. Implica preguntar por la frecuencia con que se leen los siguientes formatos: diarios y periódicos, revistas, cómics y tebeos, libros y otros.

— Género de las lecturas. Se refiere al contenido que prefieren leer en los libros; implica preguntar por la frecuencia con que leen libros de: humor, aventuras y fantasía, ciencia ficción, divulgación histórica y científica, romanticismo y otros.

— Orden de preferencia de los temas que leen en los materiales impresos cualquiera que sea su formato. Implica preguntar la ordenación de más a menos importante de los temas que les gustan leer en revistas, prensa, cómics, libros y otras publicaciones.

— Nombre y género literario del último libro leído.

3) EJEMPLO PARCIAL DE DISEÑO DE LA CÉDULA DE UNA ENTREVISTA ESTRUCTURADA

NOTA: A partir del marco elaborado se presenta un ejemplo de cédula de entrevista El ejemplo **sólo se ha desarrollado para la dimensión «preferencias literarias»,** por lo tanto, para que pueda considerarse completo, sería necesario desarrollar preguntas para el resto de las dimensiones apuntadas en el marco teórico del instrumento.

CÉDULA DE LA ENTREVISTA PARA LA EVALUACIÓN DE HÁBITOS LECTORES

EN MAYÚSCULA: INSTRUCCIONES PARA EL ENTREVISTADOR
EN MINÚSCULA: TEXTO A LEER AL ENTREVISTADO

COMENZAR CON LA SIGUIENTE PREGUNTA:

Buenos días (O, SI ES POR LA TARDE, buenas tardes). Has sido elegido para participar en un estudio de evaluación de los hábitos lectores de chicos y chicas adolescentes como tú. Me gustaría hacerte una entrevista anónima sobre este tema que durará pocos minutos. ¿Deseas participar?

SI EL ENTREVISTADO RESPONDE «NO», DECIR:

Te aseguro que la información que te solicitamos será tratada confidencialmente y la entrevista no te quitará mucho tiempo. ¿Deseas participar?

SI EL ENTREVISTADO SIGUE RESPONDIENDO «NO»:

Entonces, muchas gracias por tu atención y perdona las molestias. FINALIZA LA ENTREVISTA CON ESTE CANDIDATO.

SI EL ENTREVISTADO RESPONDE «SÍ»:

Gracias por tu participación. Como te he indicado, la entrevista no te llevará mucho tiempo y tu colaboración nos permitirá conocer cuáles son las costumbres y hábitos que los chicos y chicas de tu edad tienen respecto a la lectura. Este estudio se está realizando en la Universidad de Sevilla y de sus resultados derivarán propuestas de actuación para mejorar los servicios actualmente existentes (como los que prestan las bibliotecas públicas, o los talleres de animación a la lectura, etc.). Puedo asegurarte que la información que aportes será tratada de forma anónima y confidencial. Es por ello que te pido que seas muy sincero a la hora de contestar. Todas las preguntas que te formularé se refieren a cosas que sueles o no hacer. Contesta según tu opinión y experiencia, indicando la respuesta más adecuada para cada caso. En aquellas preguntas en las que se te ofrezcan varias alternativas de respuesta, se trata de que elijas sólo una, la que en ese caso sea más importante o frecuente para ti. Si tienes alguna duda, pregúntame antes de responder. ¿Has comprendido bien las instrucciones?, ¿quieres preguntar algo antes de empezar?

SI EL ENTREVISTADO RESPONDE QUE SÍ Y PREGUNTA CÓMO PUEDE CONOCER LOS RESULTADOS:

Al final te entregaré una tarjeta en la que figura un teléfono. En un par de meses, podremos ofrecerte un resumen de los resultados si es de tu interés.

SI EL ENTREVISTADO PREGUNTA ALGO RELACIONADO CON LAS INSTRUCCIONES YA EXPLICADAS:

Volveré a leerte exactamente la información que has solicitado (VOLVER A LEER LA PARTE DE LAS INSTRUCCIONES A LA QUE SE REFIERE LA PREGUNTA Y DECIR: ¿Te queda ahora más claro?).

SI EL ENTREVISTADO RESPONDE QUE HA ENTENDIDO TODO:

Bien, en ese caso, comenzaré a preguntarte.

ANOTAR AQUÍ EL SEXO DEL ENCUESTADO MARCANDO CON UNA CRUZ: CHICO CHICA

Las primeras preguntas que te voy a hacer se refieren a algunas características personales tuyas. ¿Puedes decirme tu edad?

EDAD: ____ AÑOS.

Bien, ¿en qué curso estás? _____ (ANOTAR CURSO y ETAPA).

¿Cuál es el nombre del colegio o centro en el que estudias?

COLEGIO: _____

¿Puedes decirme cuál es la profesión de tu padre? SI EL ENTREVISTADO NO ENTIENDE LA PREGUNTA: Me refiero en qué trabaja, concretamente el puesto que tiene y no tanto la empresa.

PROFESIÓN DEL PADRE: _____

SI DICE QUE NO TRABAJA O ES PENSIONISTA, INDICAR AMBAS SITUACIONES COMO PARADO O INACTIVO, SEGÚN EL CASO

¿Cuál es la profesión de tu madre?

PROFESIÓN DE LA MADRE: _____

SI DICE QUE ES AMA DE CASA, ANOTARLO ASÍ. SI DICE QUE NO TRABAJA, PREGUNTAR:

¿Es ama de casa, está parada o está jubilada? REGISTRAR: AMA DE CASA PARADA INACTIVA

EN LO DEMÁS, ANOTAR IGUAL QUE LA PREGUNTA ANTERIOR.

¿Cuántos hermanos tienes, excluyéndote tú?

Nº DE HERMANOS : _____

O sea, tus padres tienen (SUMAR UNO AL Nº APORTADO PARA CONFIRMAR) ___ hijos.

Muy bien, ahora voy a hacerte algunas cuestiones sobre el tema que antes te propuse: tus hábitos y costumbres de lectura.

(....)

12. A continuación te voy a indicar diferentes tipos de lectura. En cada uno de ellos, quiero que me digas con qué frecuencia sueles leer ese tipo de material. Las respuestas a dar pueden ser: nunca, de una a cuatro veces al año, de cinco a diez veces al año y más de diez veces al año.

Empezamos por: diarios y periódicos. Te recuerdo las opciones de respuesta: nunca, de una a cuatro veces al año, de cinco a diez veces al año y más de diez veces al año. SI PREGUNTA SI SE INCLUYEN LOS DIARIOS DEPORTIVOS, CONTESTAR «sí»:

_____ (CONSIGNAR RESPUESTA).

Muy bien, ahora haz lo mismo con las revistas. Vuelvo a recordarte las opciones de respuesta: nunca, de una a cuatro veces al año, de cinco a diez veces al año y más de diez veces al año. SI PREGUNTA SI SE INCLUYEN ALGÚN TIPO DE REVISTA DETERMINADO, DECIR «sí», SI SE TRATA DE UNA REVISTA (PUBLICACIÓN TEMÁTICA PERIÓDICA), Y «no», SI NO SE TRATA DE ESTE TIPO DE PUBLICACIÓN: _____

_____ (CONSIGNAR RESPUESTA).

Muy bien. ¿Y los cómics o tebeos? ¿Quieres que te recuerde otra vez la respuesta? SI DICE «sí», LEER DE NUEVO LAS OPCIONES. SI DICE «no»: _____

(CONSIGNAR RESPUESTA)

¿Y subrayar libros? ¿Quieres que te recuerde otra vez la respuesta? SI DICE «sí», LEER DE NUEVO LAS OPCIONES. SI DICE «no»: _____ (CONSIGNAR RESPUESTA)

¿Lees otro tipo de materiales impresos distintos a los que te he enumerado? _____ (CONSIGNAR «SI» O «NO»). EN CASO AFIRMATIVO, PREGUNTAR ¿qué tipo de material? _____ (CONSIGNAR RESPUESTA TEXTUAL). ¿Con qué frecuencia? Vuelvo a recordarte las opciones de respuesta: nunca, de una a cuatro veces al año, de cinco a diez veces al año y más de diez veces al año. _____ (CONSIGNAR RESPUESTA)

13. Muy bien, ahora pasamos a otra pregunta que seguirá un procedimiento de respuesta parecido a la que acabas de contestar. La pregunta es: ¿con qué frecuencia lees libros de cada uno de los siguientes géneros literarios que te voy a enumerar? Las opciones de respuesta son: nunca, de uno a tres libros al año, de cuatro a ocho libros al año, más de ocho libros al año. ¿Quieres preguntar algo antes de empezar? SI RESPONDE «SÍ», PUEDE PREGUNTAR QUÉ ES UN GÉNERO LITERARIO, EN ESE CASO, DECIR Se refiere al tema principal del libro, si es de amor, de aventuras, de guerra, etc. SI PREGUNTA SI DEBE INDICAR LA FRECUENCIA EN CADA TIPO, DECIR Sí, es necesario indicarla para todos los géneros que te voy a enumerar. Si hay algún tipo de libro que nunca lees, elige la opción nunca.

	Nunca	**De uno a tres libros al año**	**De cuatro a ocho libros al año**	**Más de ocho libros al año**
— Humor	☐	☐	☐	☐
— Aventura y fantasía	☐	☐	☐	☐
— Ciencia ficción	☐	☐	☐	☐
— Divulgación científica	☐	☐	☐	☐
— Novela romántica	☐	☐	☐	☐
— Poesía	☐	☐	☐	☐
— Novela histórica	☐	☐	☐	☐

¿Lees libros de otros géneros distintos a los que te he enumerado» SI RESPONDE «NO», CONSIGNAR «NUNCA» EN EL CUADRO DE ABAJO. SI RESPONDE «SÍ», CONSIGNAR EL GÉNERO QUE INDICA, RECORDARLE LAS MODALIDADES DE RESPUESTA Y CONSIGNAR LA ADECUADA.

	Nunca	**De uno a tres libros al año**	**De cuatro a ocho libros al año**	**Más de ocho libros al año**
Otros (indicar: _____)	☐	☐	☐	☐
Otros (indicar: _____)	☐	☐	☐	☐
Otros (indicar: _____)	☐	☐	☐	☐
Otros (indicar: _____)	☐	☐	☐	☐

14. A continuación se presentan un conjunto de temas. Ordénalos según los prefieras a la hora de leer (en libros, revistas, tebeos u otros medios), siendo 1 el que prefieres en primer lugar y 9 el que prefieras en último lugar.

Si no tienes por costumbre leer, indícamelo ahora.

Nunca leo ningún tipo de material impreso ☐

¿Has entendido la pregunta? Te la repetiré para que te quede más clara. Entre los temas que te propongo, da igual que los leas en revistas o tebeos o libros o lo que sea, tienes que indicar con un 1 el que más te gusta en primer lugar y después con un 2 el segundo que más te gusta y así hasta llegar al 9 que será el que menos te guste de todos.

REPETIR AL SUJETO LAS OPCIONES TODAS LAS VECES QUE SEAN NECESARIAS E IR CONSIGNANDO SU RESPUESTA

— Deportes Orden: ____
— Noticias políticas o sociales Orden: ____
— Noticias del corazón y de los famosos Orden: ____
— Temas de amor y romanticismo Orden: ____
— Informática y ordenadores Orden: ____
— Relatos de viajes y aventuras Orden: ____
— Temas fantásticos o del futuro Orden: ____
— Escenas de la vida real de chicos y chicas de mi edad Orden: ____
— Temas de sexo Orden: ____

15. ¿Puedes decirme ahora cuál es el título del último libro que has leído?

_____

16. Te voy a decir varios géneros literarios para que me indiques a cuál pertenece el libro que acabas de mencionar

- ☐ Humor
- ☐ Aventura y fantasía
- ☐ Ciencia ficción
- ☐ Divulgación científica
- ☐ Novela romántica
- ☐ Poesía
- ☐ Novela histórica

SI NO HA SELECCIONADO NINGUNO, DECIR ¿Te parece que pertenece a otro género distinto a los que te he mencionado o quieres que te los vuelva a repetir todos?

SI DICE QUE PERTENECE A OTRO, PREGUNTAR ¿A cuál?

- ☐ Otros (indicar: _____)

(...)

CUANDO EL ENTREVISTADO TERMINE DE RESPONDER A LA ÚLTIMA PREGUNTA, DECIR:

Bueno, aquí concluye la entrevista. Sólo me queda darte las gracias por tu colaboración. Además, como te dije al principio, te dejo esta tarjeta por si dentro de un par de meses deseas recibir información sobre los resultados. Como ves figura un teléfono al que puedes llamar si tienes interés en ellos. Muchas gracias de nuevo y buenos días (O SI ES POR LA TARDE, buenas tardes).

4) EJEMPLO PARCIAL DE DISEÑO DEL FORMATO DE UN CUESTIONARIO

NOTA: Al igual que en el ejemplo de cédula de entrevista, sólo se han formulado preguntas para la dimensión desarrollada en el marco teórico («preferencias literarias»). Para que el cuestionario estuviera completo, sería necesario desarrollar el resto de dimensiones apuntadas en el marco conceptual y formular las correspondientes preguntas.

CUESTIONARIO PARA LA EVALUACIÓN DE HÁBITOS LECTORES (Mª T. PADILLA)

El cuestionario que te presentamos está elaborado por una profesora de la Universidad de Sevilla. A través de él se persigue conocer cuáles son los hábitos de lectura de los/as adolescentes sevillanos/as. La información obtenida a través del mismo será de gran utilidad para mejorar los servicios actualmente existentes (bibliotecas públicas, talleres de animación a la lectura, etc.). Es por ello que te pedimos que colabores en esta investigación. La información que proporciones será tratada siempre de forma confidencial, por lo que te rogamos que seas totalmente sincero/a a la hora de contestar. No se trata de un examen; ni tampoco existen preguntas «correctas» o «incorrectas». En todo momento, se te pregunta por cosas que sueles o no hacer. Contesta según tu opinión y experiencia, y siempre con sinceridad.

Para contestar, señala con una X la opción que consideres que mejor se ajusta a lo que tú piensas o haces. En aquellas preguntas en las que se te ofrecen varias alternativas de respuesta que te parecen válidas, tienes que elegir una y sólo una. En ese caso, la que sea más importante o frecuente para ti. Cuando ninguna de las alternativas ofrecidas te parezca adecuada, señala la opción «otras» de la pregunta e indica brevemente tu respuesta.

Si tienes interés en conocer los resultados que se han obtenido, pregunta a tu profesor/a dentro de un par de meses. Le daremos esta información a él/ella para que te la pueda comunicar.

MUCHAS GRACIAS POR TU COLABORACIÓN

Sexo: Chica Chico Edad: _____ Curso: _____

Colegio: _____

Profesión del padre: _____

Profesión de la madre: _____

Nº de hermanos (excluido/a tú): _____

(....)

12. En cada uno de los siguientes tipos de lectura indica con qué frecuencia sueles leer cada una de ellas:

	Nunca	De una a cuatro veces al año	De cinco a diez veces al año	Más de diez veces al año
— Diarios y periódicos	☐	☐	☐	☐
— Revistas	☐	☐	☐	☐
— Cómics y tebeos	☐	☐	☐	☐
— Libros	☐	☐	☐	☐
— Otros	☐	☐	☐	☐

(Indicar: _____)

13. ¿Con qué frecuencia lees libros de cada uno de los siguientes géneros literarios?

	Nunca	De una a tres veces al año	De cuatro a ocho veces al año	Más de ocho veces al año
— Humor	☐	☐	☐	☐
— Aventura y fantasía	☐	☐	☐	☐
— Ciencia ficción	☐	☐	☐	☐
— Divulgación científica	☐	☐	☐	☐
— Novela romántica	☐	☐	☐	☐
— Poesía	☐	☐	☐	☐
— Novela histórica	☐	☐	☐	☐
— Otros	☐	☐	☐	☐

(Indicar: _____)

14. A continuación se presentan un conjunto de temas. Ordénalos según los prefieras a la hora de leer (en libros, revistas, tebeos u otros medios), siendo 1 el que prefieras en primer lugar y 9 el que prefieras en último lugar. Si no tienes por costumbre leer, pon una cruz en la primera casilla.

— Nunca leo ningún tipo de material impreso ☐

— Deportes Orden: ____

— Noticias políticas o sociales Orden: ____

— Noticias del corazón y de los famosos Orden: ____

— Temas de amor y romanticismo Orden: ____

— Informática y ordenadores Orden: ____

— Relatos de viajes y aventuras Orden: ____

— Temas fantásticos o del futuro Orden: ____

— Escenas de la vida real de chicos y chicas de mi edad Orden: ____

— Sexo Orden: ____

15. ¿Cuál es el título del último libro que has leído? _____

---

16. ¿A cuál de los siguientes géneros literarios pertenece el libro que has citado?

☐ Humor
☐ Aventura y fantasía
☐ Ciencia ficción
☐ Divulgación científica
☐ Novela romántica
☐ Poesía
☐ Novela histórica
☐ Otros (indicar: _____)

(...)

Notas

[1] Nos parece interesante nombrar aquí la Ley de Ordenación General del Sistema Educativo ya que nos permite evidenciar cómo, en los contextos en los que hay una mayor tradición orientadora y una normativa legal sobre esta cuestión, existen una serie de principios y pautas a seguir marcadas por la ley. Esto implica que el diagnóstico, vinculado a la intervención orientadora, no puede desmarcarse de estos planteamientos y debe respetarlos en su aplicación práctica. En otros contextos de intervención pedagógica, si bien no existe una regulación legal al respecto, pueden servirnos de referencia aquellos aspectos de la existente para el sistema educativo reglado que nos parezcan útiles, apropiados y relevantes.

[2] El lector interesado en conocer este código deontológico puede consultar el anexo 1 de esta obra.

[3] HARRÉ, H. y DE WAELE, J.P. (1979). Autobiography as a psychological method. En G.P. Ginsburg (ed.), *Emerging strategies in social psychological research*.

[4] GALTUNG, J. (1967). *Theory and methods of social research*. New York: Columbia University Press.

[5] MATARAZZO, J.D. (1965). The interview. En B. B. Wolman (Ed.), *Handbook of clinical psychology*. New York: Mc-Graw-Hill.

[6] WU, Y. (1967). Research interview and its measurement. *Social work, 12*, 79-87.

[7] JENSEN, B. y HAYNES, S. (1986). Self-report questionnaires and inventories. En A. CIMINERO, K. CALHOUM y H. ADAMS (eds.), *Handbook of behavioral assessment*. New York: Wiley.

[8] HAYNES, S.N. (1978). *Principles of behavioral assessment*. New York: Gardner Press.

[9] ENDICOTT, J. (1981). Diagnostic interview schedule: reliability and validity. *Archives of General Psychiatry, 38*, 1300.

[10] KIWOOD, T.M. (1988). *Values in adolescent life: towards a critical description*. Unpublished Doctoral Dissertation. School of Research in Education. University of Bardford.

[11] En realidad, no existe un consenso sobre el nivel de medida que aportan los tests y las escalas de actitudes. Además de las distintas escalas de medidas de las variables, STEVENS (1951) realizó una controvertida afirmación: los instrumentos utilizados en psicología y educación son realmente escalas ordinales y no de intervalos. Esta afirmación ha suscitado una polémica que se mantiene hasta nuestros días, y que no sólo se relaciona con la posibilidad de que los tests ofrezcan o no una medida de intervalo, sino también con el tipo de operaciones matemáticas que podemos realizar con las puntuaciones que ofrecen. Así, en los diferentes ítems de que se compone un test o una escala de actitudes no podemos estar seguros de la equivalencia de las tareas que implican (THORNDIKE y HAGEN, 1978). Por ejemplo, cómo podemos saber que un problema de cálculo (sumar 23 y 25) es exactamente igual a otro (restar 35 y 22). Si bien somos conscientes de las serias dificultades para considerar como de intervalo el nivel de medida que nos permiten los tests, nos mostramos de acuerdo con BORGATTA y BOHRNSTEDT (1981), quienes afirman que estos instrumentos son realmente *escalas de intervalo* aunque *imperfectas*; en su opinión, los tests tienen más de escala de intervalo que de escalas puramente ordinales, pudiendo, por tanto, considerarse como *escalas de intervalo con un margen de error tolerable*. Este punto de vista es el que está implícito en muchos otros autores.

[12] Como ya se ha indicado, propiamente hablando, las medidas de los tests no se sustentan sobre el criterio de «homogeneidad», exigible a toda medida cuantitativa-continua, ya que los procedimientos psicométricos no permiten comprobar si todos los ítems miden una misma cosa (GARCÍA JIMÉNEZ y GARCÍA VIDAL, 1997). Sin embargo, en función de la argumentación que realizamos, nosotros entendemos que los tests son los instrumentos que más se aproximan a cumplir este criterio de homogeneidad, especialmente si su proceso de construcción responde a las características que mencionamos en el texto.

[13] En el caso de la evaluación de las actitudes se suele utilizar la denominación de *escalas de actitudes*. Por cuestiones de espacio, aquí no vamos a realizar un tratamiento específico de este tipo de instrumento tan usado en el campo de la educación, sino que lo abordaremos de forma genérica dentro de lo que denominamos como pruebas y tests no cognoscitivos o afectivos. El lector interesado en el tema encontrará un planteamiento muy exhaustivo y completo en la obra de MORALES (2000).

[14] Un análisis más exhaustivo de los diferentes tipos de pruebas estandarizadas para la evaluación del rendimiento puede encontrarse en JORNET y SUÁREZ (1996). En este trabajo los autores proponen una tipología de pruebas cuyos componentes son: pruebas como indicadores de resultados, pruebas de certificación y admisión, pruebas de dominio, pruebas de clase y pruebas individualizadas.

[15] TEA Ediciones, S.A. Fray B. Sahagún, 24, Madrid, 28036. También dispone de centros en Barcelona, Bilbao y Sevilla.

[16] Psymtec. Alonso Cano, 66, 1º 2E, 28003 Madrid.

[17] Instituto de Orientación Psicológica EOS. Avda. Reina Victoria, 8, 2º p. 28003 Madrid.

[18] Se puede consultar un ejemplo de este tipo de tabla de especificaciones en BROWN (1980:307).

[19] Algunos de estos datos pueden ser: sexo del examinado, curso, turno, etc. Si tras su aplicación la prueba va a ser procesada y analizada estadísticamente, estos datos facilitarán la realización de algunos contrastes que pueden contribuir a mejorar la información que obtenemos sobre las características del examen y la calidad del mismo.

[20] La opinión de varios autores sobre el tiempo que puede tardar un alumno en responder en una prueba de elementos de selección de respuesta tiende a coincidir con este margen. BROWN (1980) sugiere que 1 minuto por reactivo es un tiempo suficiente. Por su parte, MEHRENS y LEHMANN (1982) estiman en unos 75 segundos la cantidad de tiempo necesaria para responder a cada reactivo, y THORNDIKE y HAGEN (1978) ofrecen un amplio margen entre 30 y 100 segundos que dependerá del nivel de complejidad del ítem.

[21] En algunos tests de aptitudes y de rendimiento, las respuestas incorrectas pueden, simplemente, no puntuar o también, según el test, puntuar de forma negativa.

[22] El lector interesado en profundizar en el cálculo de los diferentes tipos de puntuaciones realizando sencillos ejercicios puede consultar la obra de LEVINE y ELZEY (1973), además de los manuales al uso de estadística descriptiva.

$$\text{Puntuación } z = \frac{\text{Punt. Directa - Media}}{\text{Desv. Típica}}$$ [23] [24]

Cuando una distribución adopta la forma de la campana de Gauss, propia de la distribución normal, las puntuaciones directas se pueden transformar en puntuaciones z normalizadas, es decir, en puntuaciones z de acuerdo con la curva normal. Para ello, se utiliza una tabla donde ya están calculados los valores de z en la curva normal y la superficie de la curva según dichos valores (tabla de la curva normal).

[25] MORALES (2000) considera que las dos publicaciones sobre el estudio de la aquiescencia más influyentes de Cronbach son: CRONBACH, L.J. (1946), Response sets and test validity. *Educational and Psychological Measurement, 6*, 475-494, y CRONBACH, L.J. (1950), Further evidence on response sets and test design. *Educational and Psychological Measurement, 10*, 3-31.

[26] HERMAN, GEARHART y ASCHBACHER (1996:34).

[27] COLE, RYAN and KICK (1995: 41-42).

[28] COLE, RYAN and KICK (1995: 40-41).

[29] COLE, RYAN and KICK (1995:43).

[30] COLE, RYAN and KICK (1995:44).

[31] En este apartado y muchos de los siguientes, se utiliza como texto generador el artículo titulado *Nosotros*, publicado en la Revista Escuela de Adultos Polígono Sur, nº 6.

[32] Algunos de los roles de la lista de KELLY (1955) son: el propio sujeto, la madre, el padre, el hermano, la hermana, la pareja, el maestro que más le influyó, el maestro que más cuestiona, la persona más feliz, la persona con más éxito, etc.

[33] Según BOTELLA y FEIXAS (1998), este procedimiento fue propuesto por primera vez por EPTIN, F.R., SUCHMAN, D.I. y NICKESON, C. (1971). An evalution of the elicitation procedures for personal constrcusts. *British Journal of Psychology, 62*, 513-517.

[34] La versión actualmente disponible del programa RECORD es la 3.1. El programa además tiene una versión en inglés y otra en italiano. Puede obtenerse información sobre el programa y otros materiales para trabajar la rejilla de constructos personales en la dirección www.terapiacognitiva.net.

[35] ALONSO TAPIA, J. (1997). *Orientación educativa. Teoría, evaluación e intervención.* Madrid: Síntesis.

[36] ÁLVAREZ ROJO, V. (1997). *¡Tengo que decidirme¡* Sevilla: Alfar.

[37]FLANDERS, N.A. (1977). *Análisis de la interacción didáctica.* Madrid: Anaya.

[38] MINISTERIO DE EDUCACIÓN, CULTURA y DEPORTES (2001). *Encuesta sobre hábitos lectores de la población escolar entre 15 y 16 años. Proyecto de estudio.* Madrid: Dirección General de Educación, Formación Profesional e Innovación Educativa. Centro de Investigación y Documentación Educativa.

[39] ROSINÉS, M. (1996). Una apuesta a favor del placer. *CLIJ,* 78, 82.

[40] FOUCAMBERT, J. (1989). *Cómo ser lector*. Barcelona: Laia.

[41] REAL ACADEMIA DE LA LENGUA ESPAÑOLA (1992). *Diccionario de la lengua española*. Madrid: Espasa-Calpe.

[42] MORENO, E. (2000). *Evaluación de los hábitos lectores de la juventud malagueña*. Málaga: Centro de Ediciones de la Diputación Provincial de Málaga.

[43] RODRÍGUEZ, J.M., MORENO, E. y MUÑOZ, A. (1988). Aprendizaje y ambiente: influencias del entorno familiar y escolar en el aprendizaje de la lectura (I). *Revista de Educación Especial, 3,* 21-26.

[44] ARAGÓN, M.T., GÓMEZ, M.D. y JUAN, A. (1997). La importancia de los hábitos lectores. Comunicación presentada a Edutec'97. Publicación electrónica en http://www.ice.uma.es/edutec97/edu97_c4/2-4-08.htm.

[45] MORENO, E., PADILLA, M.T. y VÉLEZ, E. (1998). La familia como contexto de adquisición de los hábitos lectores: un estudio descriptivo. En V. Llorent (ed.), *Familia y educación. Una perspectiva comparada* (pp. 363-379). Sevilla: Departamento de Teoría e Historia de la Educación y Pedagogía Social de la Universidad de Sevilla.

Índice

Créditos 5
Prólogo 10
Introducción 11
CAPÍTULO 1. INTRODUCCIÓN AL DIAGNÓSTICO Y LA EVALUACIÓN EN EDUCACIÓN 14
1. INTRODUCCIÓN AL CONCEPTO DE DIAGNÓSTICO Y EVALUACIÓN 15
2. OBJETIVOS, FUNCIONES Y ÁMBITOS DEL DIAGNÓSTICO EN EDUCACIÓN 21
3. ENFOQUES DE DIAGNÓSTICO 25
3.1. Modelos tradicionales 25
3.2. Modelos conductuales 28
3.3. Modelos derivados de la psicología cognitiva 30
4. APORTACIONES DESDE LA INTERVENCIÓN ORIENTADORA: PRINCIPIOS Y MODELOS PARA EL DIAGNÓSTICO EDUCATIVO 34
5. FASES EN EL PROCESO DE DIAGNÓSTICO 42
6. EL INFORME DE DIAGNÓSTICO 46
6.1. Cuestiones previas: formación, habilidad y ética en la elaboración del informe diagnóstico 47
6.2. Propuesta de estructura del informe 48
6.3. Recomendaciones adicionales en la elaboración de informes 50
6.4. Posibles criterios para valorar informes diagnósticos 52
7. GARANTÍAS ÉTICAS Y DEONTOLOGÍA PROFESIONAL EN EL DIAGNÓSTICO Y LA EVALUACIÓN 54
7.1. Aspectos de la profesión pedagógica a incluir en un código deontológico 55
8. CONSIDERACIONES SOBRE LA SELECCIÓN DE INSTRUMENTOS PARA EL DIAGNÓSTICO Y LA EVALUACIÓN 59
Referencias bibliográficas 62
CAPÍTULO 2. INTRODUCCIÓN AL DISEÑO Y APLICACIÓN DE PROCEDIMIENTOS OBSERVACIONALES 66
1. LA OBSERVACIÓN COMO ESTRATEGIA DIAGNÓSTICA 67
1.1. Concepto y características 67
1.2. Modalidades de observación 71

1.2.1. Observación sistematizada/no sistematizada 73
1.2.2. Observación participante/no participante 76
1.2.3. Auto-observación/hetero-observación 77
2. PREPARACIÓN DE LA OBSERVACIÓN: LA SELECCIÓN DE MUESTRAS 79
3. DISEÑO Y APLICACIÓN DE PROCEDIMIENTOS OBSERVACIONALES 86
3.1. Observación sistematizada 87
3.1.1. Aspectos a tener en cuenta en la elaboración de sistemas categoriales 88
3.1.2. Proceso de elaboración de sistemas categoriales 92
3.1.3. Tipos de sistemas categoriales 97
3.2. Observación no sistematizada 101
3.2.1. Tipos de registros de observación no sistematizada 105
3.2.2. Registros in situ 105
3.2.3. Registros a posteriori: el diario 108
3.3. Los sistemas tecnológicos en la observación sistematizada y no sistematizada 111
4. GARANTÍAS CIENTÍFICAS DE LAS TÉCNICAS Y REGISTROS OBSERVACIONALES 113
4.1. Sesgos que afectan a las técnicas y registros observacionales 113
4.2. Fiabilidad y validez de los registros observacionales 115
4.3. Criterios de calidad alternativos para los registros de observación no sistematizados 117
Referencias bibliográficas 120
CAPÍTULO 3. INTRODUCCIÓN AL DISEÑO Y APLICACIÓN DE LAS TÉCNICAS E INSTRUMENTOS DE ENCUESTA 123
1. LA ENCUESTA COMO ESTRATEGIA DIAGNÓSTICA 124
2. LA ENTREVISTA EN EL DIAGNÓSTICO Y LA EVALUACIÓN 126
2.1. Entrevistas abiertas o no estructuradas 129
2.1.1. Preparación de la entrevista no estructurada 133
2.1.2. Desarrollo y aplicación de la entrevista no estructurada 136
2.1.3. El registro de la información en la entrevista abierta 141
2.2. Entrevistas estructuradas 142
2.2.1. Diferencias entre la entrevista estructurada y el cuestionario 143
2.2.2. Preparación de la cédula 145
2.2.3. Aplicación de la entrevista estructurada y el registro de la

información
3. EL CUESTIONARIO EN LOS PROCESOS DIAGNÓSTICOS 150
3.1. Concepto y características 150
3.2. Diseño de cuestionarios y entrevistas estructuradas 152
3.2.1. Proceso de diseño y preparación de cuestionarios y entrevistas estructuradas 154
3.2.2. Las preguntas en cuestionarios y entrevistas estructuradas: tipos, selección y formulación 158
3.2.3. Tendencias de respuesta que inducen al sesgo 166
3.2.4. Preparación del formato del cuestionario 169
4. GARANTÍAS CIENTÍFICAS DE LAS TÉCNICAS E INSTRUMENTOS BASADOS EN LA ENCUESTA 173
4.1. Cuestionarios 174
4.2. Entrevistas estructuradas 175
4.3. Entrevistas abiertas 179
Referencias bibliográficas 182
CAPÍTULO 4. INTRODUCCIÓN AL DISEÑO Y APLICACIÓN DE TESTS Y PRUEBAS OBJETIVAS 185
1. LA MEDIDA EN EL DIAGNÓSTICO Y LA EVALUACIÓN 186
1.1. La medición, la evaluación y el diagnóstico 192
2. LOS TESTS: CONCEPTO Y CARACTERÍSTICAS 194
3. TIPOS DE TESTS 201
3.1. Algunas clasificaciones de tests aplicables al campo educativo 202
3.2. Clasificación temática de algunos tests y pruebas para la evaluación del desarrollo personal y profesional 209
3.2.1. Pruebas estandarizadas para la evaluación de los procesos de estudio y aprendizaje 209
3.2.2. Pruebas para la evaluación de los intereses profesionales 210
3.2.3. Instrumentos para la evaluación del desarrollo vocacional y profesional y la toma de decisiones 210
3.2.4. Instrumentos para la evaluación de los valores personales y laborales 211
3.2.5. Instrumentos para la evaluación de ambientes, situaciones y problemas personales y profesionales 211
3.2.6. Instrumentos para la evaluación del autoconcepto y la autoestima 212
3.2.7. Instrumentos para la evaluación de las habilidades sociales y el autocontrol 212

autocontrol 212
3.2.8. Instrumentos complementarios en el proceso de orientación personal y profesional 212
4. CONSTRUCCIÓN DE PRUEBAS OBJETIVAS 215
4.1. Construcción de los planos de la prueba: objetivos, contenidos y tabla de especificaciones 216
4.2. Formulación y selección de elementos 224
4.3. Ordenación de los elementos y preparación del formato de la prueba 231
4.4. Valoración y puntuación 232
5. APLICACIÓN DE TESTS: PRECAUCIONES Y COMPETENCIAS 234
6. INTERPRETACIÓN DE LAS PUNTUACIONES DE LOS TESTS 238
6.1. Puntuaciones percentiles 241
6.2. Puntuaciones típicas o z y puntuaciones típicas derivadas 244
7. GARANTÍAS TÉCNICAS Y CIENTÍFICAS EN LOS INSTRUMENTOS DE MEDIDA 249
7.1. Tendencias de respuesta que afectan a la medida 249
7.2. Análisis de ítems 252
7.3. Características técnicas de los instrumentos de medida 253
7.3.1. Fiabilidad 255
7.3.2. Validez 258
Referencias bibliográficas 263
CAPÍTULO 5. OTRAS TÉCNICAS Y ESTRATEGIAS PARA EL DIAGNÓSTICO EN EDUCACIÓN 267
1. INTRODUCCIÓN 268
2. EL PORTAFOLIOS 270
2.1. Portafolios de aprendizaje 271
2.2. Propósitos de los portafolios de aprendizaje 272
2.3. Las metas de aprendizaje 273
2.4. Contenidos del portafolios 274
2.5. Las tareas de aprendizaje 277
2.6. Modelos de portafolios de aprendizaje 277
2.7. Materiales de los portafolios de aprendizaje 283
2.8. Criterios de evaluación 284
2.9. Validez de los portafolios 285
3. LA HISTORIA DE VIDA 288

3.2. La historia de vida como instrumento y método de diagnóstico e intervención educativa. Presupuestos fundamentales 291
3.3. Operativización de la historia de vida. Análisis de un ejemplo: Trozos de mi vida 296
4. EL GRUPO DE DISCUSIÓN 300
4.1. El proceso de preparación y realización del grupo de discusión 302
4.2. El rol del moderador del grupo de discusión 307
4.3. Consideraciones finales 308
5. LAS REJILLAS DE CONSTRUCTOS PERSONALES 310
5.1. Preparación, aplicación e interpretación de la técnica de rejilla de constructos personales 312
5.2. Un ejemplo de rejilla de constructos personales aplicada a la decisión vocacional 318
Referencias bibliográficas 325
ANEXOS 328
Anexo 1. CÓDIGO DEONTOLÓGICO PROFESIONAL 329
Anexo 2. PROPUESTA DE NORMAS ÉTICASPARA LA ELABORACIÓN DE UN CÓDIGO DEONTOLÓGICO DEL PROFESIONAL DE LA PEDAGOGÍA (Versión 2001) 331
Anexo 3. EJEMPLOS DE SISTEMAS CATEGORIALES DE OBSERVACIÓN 335
Anexo 4. EJEMPLOS DE APLICACIONESDE LAS TÉCNICAS DE ENCUESTA 339
Notas 357